Statistical Adjustment

for

Continuous Timber Attributes

PREPARED FOR:
British Columbia Ministry of Forests
Resources Inventory Branch
Victoria, B.C.

PREPARED BY:
Dr. Margaret Penner
Forest Biometrician
Box 407, R.R. 4
Huntsville, ON P1H 2J6

Revised November 2001

106-06472/

Executive Summary

The adjustment for category 1 variables within the Vegetation Resources Inventory was examined. The algorithm consists of using unbiased ground sample (Phase II) information to adjust the possibly biased photo interpretation (Phase I) estimates. The adjustment is based on the Phase I estimates and the relationship between the ground and the photo estimates.

The species composition adjustment is addressed in a separate report and three versions were implemented here. The first version, adjustment by leading species (ADJ_{LS}), adjusted the Phase I polygon attributes based on the Phase II sample by leading species groups and maintained the Phase I species composition. The second version, adjustment with constraints (ADJ_{CONST}), first computed the population total for each attribute regardless of species using a regression estimator. Then the polygon attributes were adjusted by leading species group and constrained to sum to the computed population values. The third version, adjustment with sequential balancing (ADJ_{SQB}), adjusted the attributes by species group using sequential balancing. The first two algorithms retained the Phase I species composition. ADJ_{SQB} adjusts the species composition by adjusting each species group independently and varying the relative amounts of the species identified in a polygon in Phase I. This does not directly address the problem of species misclassification.

For ADJ_{LS} and ADJ_{CONST}, unbiased estimates are obtained by leading species strata. For ADJ_{SQB}, unbiased population totals are obtained by species group based on the relationships between Phase I and Phase II data and the Phase I population estimates. For ADJ_{CONST}, the polygon level attributes are adjusted by leading species with the constraint that they sum to the population totals. For ADJ_{SQB}, the attributes available for all species within a polygon (basal area and volume) are adjusted by species group (sp0). There is no constraint at the polygon level but the sp0 adjustments must sum to the population totals. For ADJ_{SQB}, attributes available for only the leading and secondary species (age, height, site index), have much smaller sample sizes and some species have no data. For these attributes a separate adjustment algorithm is described.

Some groups (leading species in the cases of ADJ_{LS} and ADJ_{CONST} and sp0 in the case of ADJ_{SQB}) did not have enough data to estimate a reasonable adjustment relationship (evaluated on the basis of the magnitude of the correlation and the matching success between Phase I and Phase II). These were combined into a separate group labeled "other" and a separate adjustment relationship estimated for that group.

The Phase II sampling concentrates on trees >4cm dbh. The lack of Phase II basal area and volume estimates for young stands causes some problems. For this report, the population of interest was defined as the vegetated treed polygons with a Phase I age of at least 20 years. This reduced the problem of the small trees.

A number of attribute combinations have predetermined relationships with one another (e.g., basal area = stems/ha × $\text{diameter}^2_q \times 0.00007854$ and SI = f(age, top height)). These relationships are nonlinear and it is not necessary that these relationships be maintained at the polygon and population level.

The adjustments significantly affected all attributes except top height. Significant increases in basal area and volume between Phase I and Phase II emphasized the need for adjustment. The overall increases dominated shifts in basal area by leading species or species group. Thus ADJ_{LS} and ADJ_{CONST} address the main bias in Phase I at the polygon level (all species combined) while ADJ_{SQB} provides further refinements at the sp0 level at the cost of additional computation complexity.

1. INTRODUCTION 1
2. OBJECTIVE 1
3. BACKGROUND 1
3.1 THE PROBLEM 1
3.2 THE SEQUENTIAL BALANCING ALGORITHM 2
3.3 THE DATA 2
3.3.1 Phase I 2
3.3.2 Phase 2 3
3.3.3 Differences between Phase I and Phase II attributes 3
3.3.4 Layers 3
3.3.5 Species groupings (sp0) 4
3.4 CHOICE OF ESTIMATORS 5
3.4.1 Inflation estimator 5
3.4.2 Ratio of means 5
3.4.3 Regression estimator 6
3.5 ERROR SOURCES 6
3.5.1 Misclassification error 6
3.5.2 Sampling error 6
3.5.3 Measurement error 6
3.5.4 Adjustment error 6
4. THE ADJUSTMENT 8
4.1 ADJ_{LS} 8
4.1.1 Basal area 9
4.1.2 Density 10
4.1.3 Age 10
4.1.4 Height 10
4.1.5 Site index 10
4.1.6 Consistency between adjusted age, height, and site index 11
4.1.7 Volume 11
4.2 ADJ_{CONST} 12
4.2.1 Basal area 12
4.2.2 Density 14
4.2.3 Age 14
4.2.4 Height 14
4.2.5 Site index 14
4.2.6 Volume 14
4.3 ADJ_{SQB} 14
4.3.1 Basal area 16
4.3.2 Density 16
4.3.3 Height 16
4.3.4 Age 19
4.3.5 Volume 19
4.3.6 Minor species 19
4.3.7 Phase I species or the adjusted species 19
4.3.8 Missing data 19
4.4 SPECIES ASSIGNMENTS 20
5. AN EXAMPLE 20
5.1 ADJ_{LS} 20
5.1.1 Basal area 20
5.1.2 Density 21
5.1.3 Age 22

5.1.4 Height 22
5.1.5 Volume 22
5.2 ADJ_{CONST} 23
5.2.1 Basal area 23
5.2.2 Density 24
5.2.3 Age 24
5.2.4 Height 24
5.2.5 Volume 25
5.3 ADJ_{SQB} 25
5.3.1 Basal area 25
5.3.2 Density 26
5.3.3 Age, height, and volume 26

6. DISCUSSION 27
6.1 DEFINING THE POPULATION 27
6.2 COMPARISON BETWEEN ALGORITHMS 28
6.3 PROPERTIES OF THE ESTIMATES 30
6.3.1 Appropriate estimators 30
6.3.2 Bias of Phase I estimates 30
6.3.3 Population and stratum estimates 31
6.3.4 Polygon estimates 33
6.4 SPECIES MISCLASSIFICATION 33
6.5 SPECIES LABELS 34
6.6 COMPATIBILITY AMONG VARIOUS UTILIZATION LIMITS 34
6.7 COMPATIBILITY BETWEEN ATTRIBUTES 34
6.8 YOUNG STANDS 34
6.9 INCORPORATION OF THE ADJUSTMENT INTO THE DATABASE 34

7. OUTSTANDING ISSUES 35
7.1 POLYGONS WITH NO BASAL AREA 35
7.2 LESS COMMON SPECIES 35
7.3 PLAUSIBLE ESTIMATES 36
7.4 CUSTOMIZATION OR LOCALIZATION OF THE ADJUSTMENT 36

8. CONCLUSIONS 36

9. LITERATURE CITED 37

10. APPPENDIX A: AN EXAMPLE OF THE VARIOUS ADJUSTMENT ALGORITHMS 39
10.1 THE DATA 39
10.2 THE ALGORITHMS 40
10.2.1 ADJ_{LS} 40
10.2.2 ADJ_{CONST} 41
10.2.3 ADJ_{SQB} 42

11. APPENDIX B: VRI ADJUSTMENT ATTRIBUTES LIST (FROM B.C. 1999B) 44

1. Introduction

The Vegetation Resources Inventory (VRI) undertaken by the Resources Inventory Branch of the B.C. Ministry of Forests is designed to provide good estimates of the forest estate for its main clients: Timber Supply Branch, the forest districts and the regions. The data collection has two phases. Phase I is the photo interpretation phase which consists of polygon delineation and estimation of many polygon parameters including area, species composition, basal area (BA), trees per hectare (TPH), age, height, and site index (SI). Phase II is the ground sampling phase and consists of placing ground sample plots (each consisting of a 5 plot cluster) in a subset of the Phase I polygons and estimating the same variables (except area) as well as the size distributions and various volumes. These two phases are each compiled to describe the polygons.

The VRI adjustment consists of using the unbiased Phase II estimates and the biased Phase I census to produce unbiased inventory unit estimates. These unbiased estimates are then distributed to the polygons. Although the adjustment process is often thought of as "adjusting the Phase I estimates" it also includes using the Phase II data to adjust some attributes not estimated in Phase I, notably volume.

To be current, the VRI must allow for updates in terms of growth and depletions. The growth model for the VRI is VDYP7 (still under construction). Therefore, although the clients specify the outputs of the VRI, some additional variables may be required by VDYP7 in order to provide the clients with updates.

This report outlines general VRI adjustment procedures. The Sunshine Coast District was used as an operational pilot study to demonstrate the techniques and some of the complications.

Three algorithms are presented 1) adjustment by leading species group (ADJ_{LS}), 2) adjustment by leading species group subject to an overall constraint (ADJ_{CONST}), and 3) sequential balancing (ADJ_{SQB}). Although the first two algorithms do not satisfy the requirement of unbiased species level estimates for the population, they have historic interest and are presented here for comparison.

The VRI Adjustments Working Group (B.C. 1999b) identified category 1 variables requiring adjustment (see Appendix B).

2. Objective

The objectives of this project were to

- explore methods for generating derived attributes;
- develop adjustment methodologies for all attributes in Category 1 (see Appendix B);
- make recommendations on a process for choosing between options for adjustment methodologies;
- write detailed procedures implementing the recommended methodologies; and
- report on the results.

The working data set was from the Sunshine Coast District. This is an operational data set. The coefficients and statistics presented throughout the report were based on the Phase I and Phase II data for the Sunshine Coast District. Coefficients would have to be estimated separately for other inventory units.

3. Background

3.1 The problem

The problem of inventory adjustment can be stated as follows:

> **How should the Phase I data be used to derive population estimates from the Phase II sample?**
>
> **How should these population estimates then be distributed back to the polygon level?**

The first problem is relatively straightforward. Using basal area as an example, good population estimates can be derived for all species combined and by species.

Allocating these population estimates back to the polygon level is neither a simple nor a trivial problem. Again, using basal area as an example, the Phase I polygon estimates of basal area of balsam fir may be incorrect because the polygon basal area was incorrect. It could also be incorrect because the species composition was incorrect (but the polygon basal area was correct). In practice, both sources of error are present. For instance, in the Sunshine Coast District data set used in this report, in the 102 ground sampled polygons, Phase I generally underestimated the polygon basal area. As well, of the 102 ground sampled polygons, only one had a Phase II basal area of zero while 14 had a Phase I basal area of zero. The same two sources of error occur at the species level. The average white pine basal area, when it occurs in the ground sample, is approximately $4m^2/ha$ in both Phase I and Phase II. However, it occurs in two ground sampled polygons in Phase I and in seven polygons in Phase II. To match the ground sample data, white pine needs to be added to some Phase I polygons. Cedar occurs in more polygons in Phase II and in greater amounts.

Consistent biases in polygon basal area estimation can be reduced using regression and ratio estimators. Inaccuracies in species composition require some method of adding and deleting species from polygons or moving species from one polygon to another.

This report focuses on adjusting over and underestimation errors, not on species misclassification. However, the topic of species misclassification is discussed.

3.2 The sequential balancing algorithm

Sequential balancing was developed to adjust species composition and has been described by Iles (1996) and J.S. Thrower and Associates (1998). Penner (2000) prepared a report on the adjustment of species composition using sequential balancing and recommended using basal area and some sort of species grouping (sp0 or genera) for the adjustment. The algorithm described in that report was the algorithm used as ADJ_{SQB} in this project. The algorithm outputs adjusted basal area, by species, constrained to sum to the unbiased adjusted population total and reflecting the species composition of the ground sampling.

3.3 The data

The working data for this project came from the Sunshine Coast District.

3.3.1 Phase I

The TSA portion of the Sunshine Coast District data set consists of 40,361 polygons, representing a total area of 1,139,039 ha. The data were updated to 1998 (the year of the Phase II sampling) and came from a variety of sources (not necessarily new aerial photography). Only the vegetated treed polygons (polygons in which the B.C. land class code begins with "VT") were considered further here. Four polygons were designated as "interior" but none were vegetated or had any volume. All volumes were compiled to the coastal primary utilization level of 17.5cm.

Basal area was not available from Phase I so it was estimated for this study. Basal area should be available from Phase I in future inventories. Basal area (BA) was estimated using the following equation.

$$BA(m^2/ha) = \frac{1}{0.3} \cdot \frac{volume\,(m^3/ha)}{dominant\ height\,(m)}$$

Species basal area was estimated as the total basal area multiplied by the species composition percentage divided by 100.

Density (trees per hectare) was available for less than 20% of the Phase I polygons with a positive basal area. It too should be available from future inventories (B.C. Ministry of Forests 1999b) so missing values were estimated by

density = 398.697851 + 48.070305*BA - 3.049674*vol - 2.88217*agesp1

if volume and age were present or by

density = 611.681676 + 8.205441*BA

if volume or age were absent or the density prediction was less than 50 stems/ha.

It was not the purpose of this investigation to develop methods to complete phase I data. The equation used here to estimate basal area is a reasonable approximation and the equation for density is a regression estimator based on the available data. These are applicable only to the Sunshine Coast District and may not be the most appropriate equations for this data set or any other data set.

For reasons given in section 4.1.7, polygons with a Phase I age less than 20 years were not included in the adjustment. This removed less than 4% of the polygons (Table 1).

Table 1. The polygons in the Sunshine Coast District are summarized by age category.

Phase I age	Number of polygons	Area (ha)	Average Phase I basal area (m^2/ha)	Average Phase I height (m)
< 20 years	1,504	21,609	0.099	3.502
≥ 20 years	32,377	594,718	33.340	26.489

3.3.2 Phase 2

The Phase II sampling took place in 1998 and consisted of two sets of field samples. Cluster IDs DSC1-0023-O, DSC1-0025-O, DSC1-0032-O, DSC2-0004-Z, DSC2-0006-Z, and DSC2-084-Z were not included in the analysis. Due to the needs in the area, two strata were identified – operable and inoperable, each with a different sampling intensity. Therefore, a weight variable was assigned to each ground sample representing the relative selection probability of that ground sample (see Table 2 - 2.716 = 264436/24 +348891/86).

Table 2. The Sunshine Coast District is summarized by operability code and the weight associated with each ground sample.

Operability Code	Description	Area (ha)	Ground samples	Weight
None, I, N	inoperable	264,436	24	2.716
A, C	operable	348,891	86	1.0

3.3.3 Differences between Phase I and Phase II attributes

There were a number of differences between the Phase I and Phase II attributes. The Phase I attributes generally reflected the entire polygon while the Phase II ground plots sampled only a portion of the polygon. The Phase I attributes may have been recent photo estimates or may have been based on historic data that was updated. Both the original data and the updating process may have included errors. The Phase I attribute definitions may have been different from Phase II. For example, the current Phase I procedures estimated the leading species age as the average age, weighted by basal area, of the dominant, codominant, and high intermediate trees of that species. The Phase II age was the average breast height age (unweighted) plus an age to breast height estimate.

The Phase I leading species may have been different than the Phase II leading species resulting in different estimates of leading species age, height, and SI. In the subsequent algorithms, the Phase I attributes were used as the basis of stratification.

All these differences added uncertainty to the Phase I-Phase II adjustment relationships.

3.3.4 Layers

The photo interpretation guidelines (B.C. MoF 1999a) give the following description of layers:

> Tree layers are distinguished according to recognized height differences that are, in many cases, associated with distinct age differences. Identification guidelines may vary, depending on each user's needs and the complexity of the project layer. An example of this is a regenerated lodgepole pine stand growing under an older Douglas-fir layer after a fire.

In the Sunshine Coast District, the layers in Table 3 were identified.

Table 3. The phase I data for the Sunshine Coast District is summarized by layer.

Layer	Interpretation	Number of polygons	Area (ha)
1	Primary layer	33,881	613,327
2	Secondary layer	1,001	17,839
S	Silvicultural layer	846	13,283
V	Veteran layer	1,599	30,053

The identification of layers during ground sampling is problematic. *In the case of a veteran layer, the ground sample is not extensive enough to include more than a couple of older larger trees making it difficult to establish whether these represent a true polygon layer or just a couple of big trees.* There has been some discussion on the problem of ground sample layers with regards to leading species. The leading species is generally considered to be the dominant species with regards to management planning. *Although a veteran layer may dominate in terms of basal area, it is not generally the layer being managed so the leading species should not be determined from the veteran layer.* As a result of this problem, there have been some attempts to define the leading species. Essentially, if the total basal area/ha at the 17.5 cm+ utilization level is not at least 50% of the total basal area/ha at the 4 cm+ utilization level, it is assumed there is a veteran layer and it is dominating the species composition at the 17.5cm+ utilization. Therefore, to include the primary layer species, the utilization level is lowered to 12.5cm. Again, if the total basal area/ha at the 12.5cm utilization level is not at least 50% of the total basal area/ha at the 4 cm+ utilization level the utilization level is lowered to 7.5cm. This lowering of the utilization level is continued until the basal area is at least 50% of the 4cm+ basal area. At this final utilization level, the leading species is determined. Note that this lowering of the utilization level is only for the purposes of determining the leading species. All polygons were adjusted at the same utilization level.

The current plan of the Ministry of Forests does not involve adjusting the veteran layer. The 'S' layer represents what silviculturalists would like to see, not what is actually present. On the recommendation of Cam Bartram, these were removed from the phase I data prior to adjustment.

It was beyond the scope of this project to develop an algorithm for determining phase II layers purely from the data. The photo interpretation guidelines for the amalgamation of layers (BCMOF 1999, section 11.6) combine the layers for a given polygon by summing the basal area and the trees per hectare. The species composition is calculated as the proportion of the total basal area (all layers combined) for a given species. The leading species height and age to be adjusted is calculated as the average of the layers, weighted by the layer basal area. The change in height and age as a result of adjustment is then reallocated to the layers in proportion to the layer basal area.

VDYP6, unlike VDYP7, does not estimate the volume of the veteran layer. This greatly reduced the usefulness of the data and the correlations with the ground sampling. This layer had low crown closure (5% or less) but the ground sampling certainly indicated the presence of some volume. There are two options for estimating the volume of this layer: 1) run the veteran layer through VDYP6 separately, specifying the data correspond to the primary layer or 2) derive an empirical relationship between age, height, site index, and stocking for each leading species in the database.

Only the primary layer was included in the adjustment.

3.3.5 Species groupings (sp0)

The Ministry of Forests has used 16 species groups (Table 4) to compile volume from growth and yield samples. These are not necessarily applicable to other inventory uses. The groups are used by VDYP7.

In general, the phase II sampling identified trees to the species level while in phase I the trees were sometimes identified to the species level (e.g., HW) and sometimes to the genus level (e.g., H). Due to limited sampled sizes, it may be necessary to group some sp0s together for the purposes of adjustment. The individual sp0 label is retained but a group adjustment relationship is used.

Table 4. The species associated with the 16 sp0 codes are given.

Sp0	Genus	Species codes	Interpretation
AC	A	Anything but AT starting with 'A'	Cottonwood
AT	A	AT	Aspen
B	B	Anything starting with 'B'	Balsam (abies)
C	C	Anything starting with 'C' (except 'CP' or 'CY'), 'I', or 'J'	Cedar (excluding yellow cedar)
D	D	XH, anything starting with 'D', 'V', or 'W'	alder
E	E	Anything starting with 'E'	birch
F	F	XC, anything starting with 'F'	Douglas-fir
H	H	TW, anything starting with 'H'	Hemlock
L	L	Anything starting with 'L'	larch
MB	M	GP, QG, RA, anything starting with 'K', or 'M'	Maple
PA	P	PA, PF	White-bark pine
PL	P	Anything but PA, PF, PS, PW, or PY starting with 'P'	Lodgepole pine, and jack pine
PW	P	PS, PW	White pine
PY	P	PY	Yellow pine
S	S	Anything starting with 'S'	Spruce
Y	Y	CP, CY, anything starting with 'Y'	Yellow cedar
blank	blank	Everything else	invalid code

3.4 Choice of estimators

The algorithms described here use estimators from the family of regression estimators – the ratio of means estimator and the linear regression estimator. Each estimator has its own properties. The most appropriate estimator for a given situation depends on the data and the whether certain assumptions are appropriate.

All the estimators used here assume a linear relationship between the Phase I and Phase II attributes.

Although each of the algorithms described here is implemented with a specific estimator, the data should be examined and the reasonableness of the assumptions associated with each estimator examined. The estimator used in the example was not always the most appropriate for all the variables.

3.4.1 Inflation estimator

Although the inflation estimator is not a regression estimator and was not used in the algorithms, in some cases it is a better choice than the regression estimators. The inflation estimator simply uses the Phase II data to estimate the population parameters. It does not require any assumptions about the relationship between x and y. It may be the lowest variance estimator when the correlation between x and y is very low. Note the correlation only measures the linear association between x and y. The association may be non-linear of the x variable may require transformation to improve the relationship.

3.4.2 Ratio of means

The ratio of means estimator provides unbiased estimates when the relationship between x and y is linear and through the origin. If the variance of y for a given x is proportional to x, the ratio of means estimator is the best (lowest variance) linear unbiased estimate.

The ratio of means estimator is often appropriate for situations such as the inventory adjustment where both x and y are estimates or measurements of the same quantity. It is often reasonable to assume that if x=0, y=0. In addition, with many biological attributes, the variance increases with the magnitude of the attributes.

If the Phase I estimates have a proportional bias (e.g., the tree heights are, on average, over-estimated by 10%), the ratio of means estimator will provide unbiased estimates. If, in addition, the variance of the Phase II measurements is proportional to the Phase I estimates, the ratio of means estimator is the best linear unbiased estimate.

3.4.3 Regression estimator

The regression estimator provides unbiased estimates when the relationship between x and y is linear. If the variance of y for a given x is constant the regression estimator is the best (lowest variance) linear unbiased estimate.

If the bias of the Phase I estimates has a constant component as well as a proportional component (e.g., the densities are, on average, over-estimated by 10% and the first 100 trees are always ignored), the regression estimator will provide unbiased estimates. If, in addition, the variance of the Phase II measurements is constant relative to the Phase I estimates, the regression estimator is the best linear unbiased estimate.

It can be shown that the variance of the inflation estimator of the mean will be less than the variance of the regression estimator if the sample correlation $r^2 < 1/(n-1)$. The reduced degrees of freedom associated with the regression estimate will also increase the width of the confidence interval slightly. Cochran (1977) points out that the superiority of the linear regression estimator is strictly a large-sample result. In small samples on natural populations, the regression estimate appears disappointing in performance. Therefore, the regression estimator does not always produce better estimates than the inflation estimator.

3.5 Error sources

There are a number of sources of error in the VRI (see Table 5).

3.5.1 Misclassification error

It is possible that some of the polygons are misclassified and are put in the wrong stratum. For instance, some of the polygons classified as non-vegetated in Phase I are found to have appreciable vegetation when ground-sampled. As well, the ADJLS and ADJCONST stratify the population based on the Phase I leading sp0. Misclassification errors are seldom addressed (but see Andersson 1994). Of the 102 ground sampled polygons, 49 had the same leading species in Phase I and II.

3.5.2 Sampling error

The sampling error refers to the errors arising from using a sample rather than measuring the entire population. In Phase II, two major sources of sampling error occur due to within polygon variation and between polygon variation. The Phase II ground cluster summaries are an estimate of the polygon attributes but there can be considerable variation within a polygon and one randomly selected ground cluster may be considerably different from another randomly selected cluster within the same polygon. Likewise, there is considerable variation between polygons and the polygon selected

Some of the sampling errors are given in this report.

3.5.3 Measurement error

The Phase II data includes measurement error and the derived variables (e.g., site index, volume) may contain estimation error. The measurement error is controlled by field sampling specifications and monitored by an audit program.

The Phase I data contains estimation error. The photo interpreters estimate some of the Phase I attributes using approximate relationships and expert knowledge. For instance, the basal area may be estimated from the species, age, height, and crown closure.

The Phase II data may also contain errors due to updating. Polygon attributes may be "grown" using simulation models. As well, the attributes for some polygons may be updated following targeted surveys such as regeneration surveys and pest impact assessments.

3.5.4 Adjustment error

Errors could arise in the adjustment as a result of using an inappropriate or inefficient model as well as using an inefficient estimator.

Appropriate model

Use of an inappropriate estimator (e.g.. a ratio estimator in the presence of a significant intercept) can result

in biased predictions. There has been considerable discussion about the use of ordinary least squares in cases like this where the independent variable is measured with error.

Benefits of stratification

All the algorithms described here are compatible with stratification. Does computing a separate regression for each stratum improve the estimates? In Table 6, the F statistic for testing the null hypothesis of a common regression slope (for all strata) is 2.12 compared to a critical value of $F_{6,94,0.05} = 2.20$. There is not much statistical evidence to support separate regression estimates for each stratum. The F-statistics for density, age, height, and volume are 0.42, 1.86, 2.50, and 3.14 respectively.

Appropriate estimator

A regression estimator does not always result in an improvement (decrease in variance) over the simple random sample estimator. In the population used here, the regression estimator performs poorly for the density adjustment.

Table 5. Some of the error sources in the inventory are identified.

Error Source	Level	Description	Quantifiable?
Misclassification error	Population Stratum	Polygons may be incorrectly included or excluded from the population (VT) or incorrectly assigned to a stratum.	Yes, by comparing the Phase I classification to the Phase II classification.
Sampling error – within polygons	Polygon	Arises from the ground cluster sampling the polygon rather than complete measurement	Yes, by specialized studies. Not quantifiable by standard VRI sampling (1 ground cluster/polygon)
Sampling error – between polygons	Stratum	Arises from sampling a fraction of the polygons rather than all the polygons. Does the sample represent the population?	Yes, can be routinely estimated
Measurement error – Phase II	Ground cluster	BA, topht, species composition, age, and density are all estimated directly but may contain measurement error. Volume and site index are derived variables and contain prediction error.	Yes, the measurement errors can be quantified using an audit. Standards (allowable errors) should be set. The compilation errors are generally unknown.
Measurement errors – Phase I errors (quantitative)	Polygon	The errors in the photo estimates contain bias and random (unexplained) variation. The Phase I estimates may also contain errors incurred by processing the photo estimates (e.g., updating to a more recent year). There are also errors in polygon delineation and assignment to strata (e.g., BEC).	The variation in photo estimates can be quantified by a study comparing photo interpreters. The bias is estimated by the adjustment. However, it is difficult to separate the individual sources of error (updating vs. photo interpreter).
Measurement errors – Phase I (qualitative)	Polygon	The dominant error is species misclassification. The photo interpreter may assign the wrong species or overlook minor species.	This can be approximated by comparing the photo estimates and the ground measurements. However, the qualitative errors are confounded by the quantitative errors.
Adjustment errors	All levels	Bias – can result from an inappropriate estimator (e.g., ratio rather than estimator) Precision – inappropriate assumptions about the variance structure can lead to inaccurate precision estimates.	Yes, by examining the adjustment statistics and, most importantly, plotting the data.

4. The Adjustment

The key variables for adjustment are basal area (by species); age, height and site index of the leading species; age, height and site index of the second species; density; and volume (by species). Collectively, these attributes plus shrub height are referred to as "Category 1" variables in the VRI Adjustment Working Group final report (B.C. 1999b, see Appendix B). Lorey height was identified by the VRI Adjustment Working Group as category 2 attribute to be adjusted while quadratic mean diameter (dbh_q) was identified as a category 2 attribute to be calculated. The distinction between category 1 and category 2 is somewhat vague but, except for volume and site index, category 1 attributes will be available from any future photo interpretation. Category 2 attributes are generally derived from category 1 attributes. "Volume" includes gross whole stem volume, net decay, net decay and waste, net decay waste & breakage, and close utilization (net top & stump). In addition, any of the attributes can be compiled to the 4.0, 7.5, 12.5, 17.5, or 22.5cm utilization levels. The utilization level is the lowest breast height diameter of trees included in the compilation.

4.1 ADJ_{LS}

The adjustment algorithm by leading species (ADJ_{LS}) is given in Figure 1 and described in the following sections. The chief feature of this adjustment algorithm is that attributes are adjusted by leading species with no constraints. The population is post-stratified into strata by leading species groups. Each stratum is adjusted independently. The Phase I species composition is maintained at the polygon level and the population values are the sum of the adjustments by leading species (LS) strata. The Phase II species composition is not used in this adjustment algorithm.

1. Adjust by leading species group stratum (h) $\hat{\bar{Y}}_h = \bar{y}_h + b_h(\bar{X}_h - \bar{x}_h)$	$\hat{\bar{Y}}$ = estimate of the average basal area per hectare (BA) for the leading species (LS) h stratum $\bar{X}$ = Phase I population area weighted mean BA for the h stratum $\bar{x}$ = Phase I sample mean BA for the h stratum $\bar{y}$ = Phase II sample mean for the h stratum Stratum membership is determined by the Phase I leading species. All polygons with a species composition are included. BA is the basal area per hectare of all species combined.
2. Compute ROM estimator by stratum $r_h = \frac{\hat{\bar{Y}}_h}{\bar{X}_h}$	r_h = ratio of means estimator for stratum h
3. Adjust the polygon BA $\hat{y}_{ih} = r_h \cdot x_{ih}$	$\hat{y}_{ih}$ = estimate of the average basal area per hectare (BA) for polygon i in the h stratum x_{ih} = Phase I basal area per hectare (BA) for polygon i in the h stratum
4. Repeat steps 1 through 3 for polygon density, age, and height.	
5. Estimate the average polygon volume/ha (all species combined) using VRIStart[1] and the adjusted polygon BA, density, age, and height.	
6. Repeat steps 1 through 3 for volume/ha.	

Figure 1. A flowchart of the ADJ_{LS} adjustment algorithm.

[1] VRIStart has not been finalized. The coefficients and estimates used here are based on the March 17, 2001 version.

4.1.1 Basal area

Basal area was computed, by leading species, using a regression estimator as follows.

$$\hat{\bar{Y}}_h = \bar{y}_h + \beta_h \cdot (\bar{X}_h - \bar{x}_h)$$

where

$\hat{\bar{Y}}_h$ = the adjusted mean for leading species stratum h

(1) $\bar{y}_h$ = the Phase II sample mean for stratum h

β_h = the least squares regression estimator based on the ground sampled polygons in stratum h

$\bar{X}_h$ = the Phase I population area weighted mean for stratum h

$\bar{x}_h$ = the Phase I sample mean for stratum h

A separate regression was estimated for all leading species groups (based on the Phase I leading species) with at least 10 ground sampled polygons. The leading species groups with less than 10 ground sampled polygons were pooled into an "Other" leading species group. A regression estimator was used to compute the "Other" leading species group as well.

The variance associated with the computed averages, valid for large samples obtained from simple random sampling (Cochran 1977 section 7.4), is as follows.

(2) $$\text{var}(\hat{\bar{Y}}) = \frac{1-f}{n \cdot (n-2)} \cdot SSE$$

SSE is the sum of the square error terms in the regression. Note that in future, the variance formulas will be changed to reflect the sampling with probability proportional to size with replacement.

The finite population correction factor (f = (N-n)/N) was ignored in this report and the following formula used.

(3) $$\text{var}(\hat{\bar{Y}}) = \frac{MSE}{n}$$

The mean squared error (*MSE*) is the *SSE* divided by the error degrees of freedom (n-2). The 95% confidence interval is as follows.

(4) $$\hat{\bar{Y}} \pm t \cdot \sqrt{\text{var}(\hat{\bar{Y}})} \approx \hat{\bar{Y}} \pm 1.96 \cdot \sqrt{\text{var}(\hat{\bar{Y}})}$$

Ratios were used to adjust the Phase I polygon estimates, by leading species group. The ratio was based on the predicted stratum mean divided by the Phase I area weighted stratum mean.

(5) $$r_h = \frac{\hat{\bar{Y}}_h}{\bar{X}_h}$$

The population totals for each leading species were estimated as follows.

(6) $$\hat{Y}_h = \hat{\bar{Y}}_h \cdot \text{area}_h = r_h \cdot \bar{X}_h \cdot area_h$$

Here, *area* is the total area associated with that leading species stratum (based on the Phase I leading species group) in the population.

The issue of young stands is discussed in sections 6.1 and 6.8 but warrants some mention here. The ground sampling has a lower dbh limit of 4cm. Trees smaller than 4cm in dbh are not measured so plots with all trees < 4cm dbh have a basal area and volume of zero. The adjusted basal areas (population, strata, and polygon) are therefore the basal areas of tree > 4cm dbh. Several options are available for dealing with the trees < 4cm dbh.

- Redefine the population so to remove trees < 4cm dbh. For instance, a minimum polygon age could be specified. The basal area of polygons younger than the minimum age would not be adjusted. Alternatively, a minimum basal area could be specified. The basal area of polygons with less basal area than the minimum would not be adjusted.
- Accept the Phase II definition of the attribute (i.e., basal area is the basal area of trees > 4cm). The adjustment is applied to all polygons regardless of age or basal area. Whatever the Phase II definition of basal area (the basal area of trees larger than 4cm dbh or the basal area of trees > 7.5cm dbh), it becomes the definition for the predictions. That is, the adjusted basal area is the basal area of all trees > 4cm dbh. The Phase I variable does not need to have the same definition as the Phase II variable (e.g., it could be the basal area of trees >5cm dbh or it could be the basal area of trees > 2cm dbh) as long as it is correlated with the Phase II variable.

For this exercise, the adjustment was limited to polygons with a minimum Phase I age for the leading species of 20 years.

4.1.2 Density

The density (average number of stems/ha, all species combined) was adjusted in a manner similar to basal area. Regression estimators were used to compute the population averages, by leading species group. The polygons were then adjusted using a ratio. The regression estimator used the relationship between the Phase I and Phase II data for the ground sampled polygons as well as the entire Phase I data set to produce an estimate combining the Phase II sample measurements with the Phase I census estimates.

The adjusted polygon densities were then input into VRIStart that then computed densities by species group.

4.1.3 Age

VDYP7.0 requires the age of the primary species and the age of the secondary species as input data (IPSJF106.DOC, Sept 1998). Age can be used in determining the site index of the primary and secondary species (also required input into VDYP7).

Ideally, leading species age should be adjusted by leading species group in a manner similar to basal area and density. First the population primary species age was computed, by leading species group, using a regression estimator. The polygons were then adjusted using a ratio estimator.

4.1.4 Height

Height was adjusted in a manner similar to age.

4.1.5 Site index

The site index associated with the leading species (and secondary species, where available) is determined as part of Phase I. The VRI photo interpretation procedures (B.C. 1999a) state the following:

> Site index estimates are required on all treed polygons, as well as polygons that are potentially capable of producing trees. Stands that are 30 years or older will have the site index derived from species, age, and height. The photo interpreter can still override the process if they:
>
> - have field data to provide site specific values; or,
> - have evidence to indicate abnormal site suppression exists.

So, for stands younger than 30 years, the photo interpreter estimates the site index and it becomes the basis for adjustment. For stands 30 years or older, there are two options for adjusting site index – 1) adjust the original Phase I site index or 2) adjust the site index estimated from the adjusted height and age. The adjustment process is designed to yield unbiased estimates so both options will lead to similar population estimates but will likely yield different estimates at the polygon level.

Generally, however, the Phase I information for young stands comes from silviculture survey data (not photo interpretation) and includes relatively accurate ht and age estimates. The update algorithms may introduce considerable noise.

Site index (SI) can be adjusted in a similar manner as age and height – by leading species group. In Phase II, the site index for the leading species (and sometimes the secondary species) is estimated as the average of the site indices of the site trees for that species, measured at the 5 plots in the VRI Plot cluster.

Adjustment of SI will bring it closer to the Phase II estimates of SI. It is not clear that the Phase II estimates of SI are a better estimate of the polygon SI than the SI based on adjusted polygon leading species age and height. As with all attributes, it is difficult to assess the goodness of the adjusted attribute since no independent estimates exist.

Recommendation – The photo-interpreted SI (polygons < 30 years old) should be adjusted if data are available. In the Sunshine Coast District, of the ground sampled polygons, 23 had a Phase I age of < 30 but only one had a non-zero age (an additional polygon had a Phase I age of 30). It appears the SI adjustment for the young polygons may be limited by the availability of data.

The merits of adjusting SI obtained from silvicultural records is questionable and should not be undertaken unless the surveys were taken many years ago and if sufficient data are available.

The derived SI (polygons ≥ 30 years old) should be based on the adjusted height and age. It does not require adjustment. If adjustment is desired, it should be based on the SI derived from the adjusted height and age (option 2 in paragraph 2 of section 4.1.5).

4.1.6 Consistency between adjusted age, height, and site index

SiteTools©, the software used for estimating SI (or any one of age, height, or SI) in Phase II, assumes a fixed relationship between age, height and SI. For a given age-height pair, only one SI is possible. However, SI is generally defined at some reference age (e.g., 50 years at breast height). If the age is that reference age, then the SI is the height. However, if the age is not the reference age, there is no guarantee the stand will follow (or has followed) the SI curve trajectory. It probably has not followed the curve exactly and it seems a bit of a waste of time to try to force the adjusted age, height, and SI to be compatible with the *estimated* relationships in SiteTools©. As well, the Phase II SI is calculated as the average SI of the sampled trees and not the SI associated with the average age and height. So there is no reason to assume that, in general, the Phase II SI should equal the SiteTools© prediction using adjusted age and Ht.

The SiteTools© predictions of site index should improve as the age approaches the reference age. The plausibility of the adjusted values increases as the estimates become closer to the SiteTools© predictions.

Recommendation – Consistency between age, height and SI is not expected nor is it required at the population level.

4.1.7 Volume

The issue of volumes of young stands is important for the Sunshine Coast District data where volume was available but basal area and density were not. In future inventories, volume will not be a Phase I attribute but will be estimated by the algorithms in VRIStart.

The whole stem volume (WSV to a 4cm utilization limit) adjustment needs to occur after the adjustment of basal area, density, and height – the variables that are required to estimate the Phase I volume. The volume adjustment should occur after the adjustment of these input variables in order to reduce consistency problems caused by adjusting all the variables independently. In this report, the Phase I volumes were estimated using the current VRIStart algorithms. These algorithms are provisional and subject to revision. These algorithms first estimate dbh_q by species as a function of basal area and density. Then Lorey height is estimated. Average tree volume is then estimated from dbh_q and Lorey height.

The adjustment of volume was similar to that of basal area. That is, the leading species strata averages were computed using a regression estimator. The polygon volumes were then adjusted using a ratio. Some leading species groups were grouped into an "other" species group and a single adjustment relationship estimated and used for all the minor species. Minor species groups were those that were rare in Phase I or absent in either Phase I or Phase II and would likely have had a poorly estimated adjustment relationship.

If the adjusted polygon species volume is implausible, there are two potential causes – the individual tree

volume or the adjusted density. Either or both of these can be readjusted in order to make the total polygon species volume more realistic. The simplest way to automate this would be to set the polygon species volume to the maximum and recalculate the density (=total polygon volume/mean tree volume). This decrease in one polygon's density would require an increase in density in one or more of the remaining polygons in order to ensure the total population density is maintained. The density of all the remaining polygons could be increased by a very small amount and each would then require a compatibility variable to maintain agreement with density, mean tree volume, and total volume. Compatibility variables are defined in IPSJF136.doc (by Jim Flewelling, dated April 10, 1999). Alternatively, a polygon with very small trees could be identified and the density increased to compensate for the polygon with an implausibly high volume. Only two polygons would require compatibility variables.

Volume, net decay waste and breakage (NDWB), is also a category 1 variable. It is estimated during the Phase II sampling but does not have an equivalent in Phase I. NDWB is always less than or equal to WSV. Two variations on ratio estimators that are used in VRIStart could be used here. One is the usual ratio estimator and the other is the use of the logit function (logit(x) = ln[x/(1-x)]). Both take the form NDWB = $\alpha \times$ WSV. The usual ratio estimator estimates α from the data with the constraint $0 \leq \alpha \leq 1$. For the logit function,

$\alpha = \text{logit}^{-1}(f(x1, x2, x3, \text{etc})$ where x1, x2, etc. could be polygon attributes such as age, DQ, etc. and $\text{logit}^{-1}(t) = \exp(t)/[1+\exp(t)]$. By definition, $0 \leq \text{logit}^{-1}(t) \leq 1$. Potentially, the logit function is more flexible (responsive to different forest conditions). However, the development of a suitable logit model is beyond the scope of this project.

4.2 Adj_{const}

The adjustment algorithm with constraints (ADJ_{const}) is given in Figure 2 and described in the following sections. The chief feature of this algorithm is adjustment by leading species group (LS) and for all species combined. Again, the population was post-stratified by LS into strata.

The adjustment for all species combined constrains the LS adjustments. The Phase I species composition at the polygon level is maintained.

4.2.1 Basal area

First, the overall basal area was computed using a regression estimator using equation (1) for all species combined.

Then, adjustment relationships were estimated by leading species group (LS) strata. Ratio of means estimators were used for the LS adjustment to avoid difficulties associated with non-zero intercepts. The LS adjustments were constrained to agree with the computed population total by multiplying the ratio by the following adjustment factor.

$$(7) \qquad k = \frac{\hat{Y}}{\sum_h \hat{Y}_h}$$

The variables are defined in step 4 of Figure 2.

Species that occurred as the leading species in less than 10 ground sampled polygons in Phase I were combined into an "other" group. An adjustment relationship was computed for this group as well. The leading species in the "other" group were adjusted to sum to the group total. For species groups found only in Phase I, the "other" adjustment relationship was used.

1. Compute the population total using a regression estimator $\hat{\bar{Y}} = \bar{y} + b(\bar{X} - \bar{x})$	$\hat{\bar{Y}}$ = estimate of the population average basal area per hectare (BA) $\bar{y} = \frac{1}{n}\sum_{n} y$ = Phase II sample mean b = least squares estimate of the regression relationship between x and y for the ground sampled polygons. $\bar{X} = \frac{1}{\sum area}\sum x \cdot area$ = Phase I population area weighted mean BA $\bar{x} = \frac{1}{n}\sum_{n} x$ = Phase I sample mean n = number of ground sampled polygons N = number of polygons in the population where the Phase I age ≥ 20.
2. compute the strata (h) ratios of means $r_h = \frac{\bar{y}_h}{\bar{x}_h}$	Stratification is by leading species group (LS) $\bar{y}_h = \frac{1}{n}\sum_{n} y_h$ = ground sample mean BA for LS h $\bar{x}_h = \frac{1}{n}\sum_{n} x_h$ = Phase I sample mean BA for LS h Ground sampled polygons are weighted by their selection probability.
3. Compute ROM estimate by stratum $\hat{\bar{Y}}_h = r_h \cdot \bar{X}_h$	$\hat{\bar{Y}}_h$ = estimated population mean BA, for LS h $\bar{X}_h = \frac{1}{N}\sum_{N} x_h$ = Phase I mean of LS h.
4. Compute the "k" adjustment factor (to force compatibility between the regression estimator for all species combined and the ratio of means estimators by LS) $k = \frac{\hat{Y}}{\sum_h \hat{Y}_h}$ $\hat{\bar{Y}}_h^{*} = k \cdot \hat{\bar{Y}}_h = k \cdot r_h \cdot \bar{X}_h$	$k = \frac{\hat{Y}}{\sum_h \hat{Y}_h} = \frac{area \cdot \hat{\bar{Y}}}{\sum_h area_h \hat{\bar{Y}}_h}$ $area_h$ = sum of polygon areas where LS= h $\hat{Y} \equiv \sum_h \hat{Y}_h$ *
5. Adjust the LS BA for each polygon $\hat{y}_{ih} = k \cdot r_h \cdot x_{ih}$	$\hat{y}_{ih}$ = estimate of the average basal area per hectare (BA) for polygon i with LS = h. x_{ih} = Phase I basal area per hectare (BA) for polygon i with LS h.
6. Repeat steps 1 - 5 for density, age, and height.	Adjust by leading sp0.
7. Estimate the average polygon volume/ha (all species combined) using VRIStart[2] and the adjusted polygon BA, density, age, and height.	
8. Repeat steps 1 - 5 for volume/ha.	Adjust by leading sp0

Figure 2. The ADJ_{const} algorithm is given.

[2] VRIStart has not been finalized. The coefficients and estimates used here are based on the March 17, 2001 version.

4.2.2 Density

The average polygon density (stems/ha, all species combined) was adjusted in a manner similar to basal area. First, the overall total was computed using a regression estimator. The regression estimator used the relationship between the Phase I and Phase II data for the ground sampled polygons as well as the entire Phase I data set to produce an estimate combining the Phase II sample measurements with the Phase I census estimates.

Then ratio of means estimators were derived by leading species and constrained to sum to the overall total. Leading species groups with low sample sizes were grouped into an "other" leading species group. These ratio of means estimators were then used to adjust the individual polygon densities.

4.2.3 Age

Leading species group age was adjusted in a manner similar to basal area. Again, if there were enough data to adjust a leading species individually, it was done. Otherwise the "other" adjustment ratio was used. "Enough data" is a subjective assessment but the higher the correlation, number of observations, and matching success between Phases I and II; the more reliable the adjustment relationship for the individual leading species group.

There is no urgent need to ensure the average leading species group age (for all polygons) is unbiased – provided the average leading species group age is unbiased. However, due to the low correlations and matching success, constraining the overall average leading species group age may be prudent.

4.2.4 Height

Height was adjusted in a manner similar to age – by leading species, grouping together minor leading species groups.

Recommendation – by default, the height adjustment should be by leading species or by order (leading or secondary) and species. It should not be just by order (regardless of species).

4.2.5 Site index

The photo-interpreted SI (polygons < 30 years old) should be adjusted if data are available.

The derived SI (polygons ≥ 30 years old) should be based on the adjusted height and age. It does not require adjustment. If adjustment is desired, it should be based on the SI derived from the adjusted height and age (option 2 in paragraph 2 of section 4.1.5).

4.2.6 Volume

The adjustment of volume was similar to that of basal area. That is, the adjustment first computed an unbiased population total for all species combined using a regression estimator. Then the species totals were computed by leading species groups (LS) using a ratio of means estimator and constrained to sum to the population total. Some leading species groups were grouped into an "other" leading species group and a single adjustment relationship estimated and used for all the minor species. Minor leading species were those that are rare in Phase I or absent in either Phase I or Phase II and would likely have a poorly estimated adjustment relationship. The individual polygons were adjusted using the same leading species group adjustment ratios.

4.3 Adj_{SQB}

The sequential balancing adjustment algorithm (ADJ_{SQB}) is given in Figure 3 and described in the following sections. The chief distinguishing feature of this algorithm is that basal area is adjusted by species group not by leading species. The population is not stratified. The basal area for each species group is estimated for the entire population. If a polygon contains F and H basal area, each is adjusted independently by a separate adjustment relationship regardless of which species is leading. The Phase I species composition is not necessarily maintained and the independent adjustments within a polygon can lead to shifts in leading species.

1. Compute the population total (all species combined) using a regression estimator $\hat{\bar{Y}} = \bar{y} + b(\bar{X} - \bar{x})$	$\hat{\bar{Y}}$ = estimate of the population average basal area per hectare (BA) $\bar{y} = \frac{1}{n}\sum_{n} y$ = Phase II sample mean b = least squares estimate of the regression relationship between x and y for the ground sampled polygons. $\bar{X} = \frac{1}{\sum area}\sum_{N} x \cdot area$ = Phase I population area weighted mean BA $\bar{x} = \frac{1}{n}\sum_{n} x$ = Phase I sample mean n = number of ground sampled polygons N = number of polygons in the population where the Phase I age ≥ 20.
2. compute the species group ratios of means for the population. $r_s = \frac{\bar{y}_s}{\bar{x}_s}$	Adjustment is by sp0 group (SG) $\bar{y}_s = \frac{1}{n}\sum_{n} y_s$ = ground sample mean BA for SG s $\bar{x}_s = \frac{1}{n}\sum_{n} x_s$ = Phase I sample mean BA for SG s Ground sampled polygons that do not contain SG s have an implied BA of 0.
3. Compute ROM estimate by species group for the population $\hat{\bar{Y}}_s = r_s \cdot \bar{X}_s$	$\hat{\bar{Y}}_s$ = estimated population mean BA, for SG s $\bar{X}_s = \frac{1}{N}\sum_{N} x_s$ = Phase I mean of SG s. Polygons that do not contain SG s have an implied BA of 0.
4. Compute the "k" adjustment factor (to force compatibility between the regression estimator for all species combined and the ratio of means estimators by species group) $k = \frac{\hat{\bar{Y}}}{\sum_{s}\hat{\bar{Y}}_s}$ $\hat{\bar{Y}}_s^{*} = k \cdot \hat{\bar{Y}}_s = k \cdot r_s \cdot \bar{X}_s$	$k = \frac{\hat{\bar{Y}}}{\sum_{s}\hat{\bar{Y}}_s} = \frac{\hat{\bar{Y}}}{\sum_{s}\hat{\bar{Y}}_s}$ $\hat{\bar{Y}} \equiv \sum_{s}\hat{\bar{Y}}_s^{*}$
5. Adjust the SG BA for each polygon $\hat{y}_{is} = k \cdot r_s \cdot x_{is}$ Note: $x_i = \sum_{s} x_{is}$	$\hat{y}_{is}$ = estimate of the average basal area per hectare (BA) for species group s in polygon i. x_{is} = Phase I basal area per hectare (BA) for SG s in polygon i.
6. Repeat steps 1 - 5 for density.	
7. Adjust height and age as in Figure 4.	
8. Estimate the average polygon volume/ha (all species combined) using VRIStart[3] and the adjusted polygon BA, density, age, and height.	
9. Repeat steps 1 - 5 for volume/ha by sp0 group.	Each sp0 within a polygon has a volume/ha estimate

Figure 3. A flowchart of the ADJ_{SQB} adjustment algorithm.

[3] The coefficients and estimates used here are based on the March 17, 2001 version of VRIStart.

4.3.1 Basal area

This step of the adjustment process adjusts basal area as well as the species composition.

The basal area was adjusted using the sequential balancing adjustment algorithm as described in Penner (2000). The inputs are the Phase I and Phase II basal areas, by species. The population basal area was calculated using a regression estimator. This was then fixed. Note that this computed population basal area was the same as that using ADJ_{CONST}. Ratio of means estimators were determined by sp0. The basal area of each sp0 was computed by polygon and the ratio of mean estimator was determined as follows.

$$r_{sp0} = \frac{\sum_{groundsample} w \cdot BA_{ph2sp0}}{\sum_{groundsample} w \cdot BA_{ph1sp0}}$$

The weighting factor (w) is included to account for the different selection probabilities for the ground sampled polygons (Table 2). The ratio of mean estimators were then multiplied by a correction factor k, similar to the k in section 4.2.1 to ensure the adjusted basal areas by sp0 summed to the computed population basal area.

The outputs were the adjusted Phase I basal areas, by species. The computed total basal area is unbiased and the total adjusted species basal areas were constrained to sum to the total of all species combined. The adjustment was by sp0 or genera. The species composition was then fixed and not altered by any of the subsequent adjustments.

The leading species may change as a result of this adjustment algorithm. In the Sunshine Coast District dataset, the main shifts in leading species were to more area in B and C and the minor species and less F and H.

4.3.2 Density

The density was adjusted as in ADJ_{CONST}

4.3.3 Height

The height of the primary species and the height of the secondary species are category I attributes to be adjusted. Height (and age) can be used in determining the site index of the primary and secondary species (also category I variables but not to be adjusted).

Discussions with Ministry of Forests staff resulted in the following assumptions which determined the adjustment algorithm. In practice, the leading species in a polygon is determined from the aerial photos and then its height estimated. The second species is then determined and its height is often estimated as a function of the leading species and the leading species height. For example, in spruce-fir mixtures, the fir height is often 80% of the spruce height. Therefore, the algorithm first adjusts leading species height by species group and then the second species height is adjusted as a function of the leading species height (i.e., the ratio of the second species height to the leading species height). In some cases, there will not be clear dominance of one species in terms of basal area and/or volume. It is anticipated in these cases there will be less agreement between the Phase I and Phase II data in terms of leading species. However, it is also anticipated that the heights will be relatively similar and the adjustments relatively similar.

In addition, missing values are allowed for height. For instance, Douglas-fir may have been identified as the leading species in Phase I and its height estimated. However, in Phase II, it may have been present (and a basal area measured) but not identified as leading or secondary and no height was measured. In Phase II, the height of Douglas-fir is a missing value. This is different from attributes such as basal area where the absence of a measurement implies a zero. For instance, if Douglas-fir was not identified on the photo, it has an implied basal area of zero.

These assumptions led to the following adjustment algorithms in Figures 4 and 5.

Note that in the Sunshine Coast District Data set, only the leading species height was available from Phase I so the second species height adjustment was not implemented.

1. Compute the population average using a regression estimator $\hat{\bar{Y}} = \bar{y} + b(\bar{X} - \bar{x})$	$\hat{\bar{Y}}$ = estimate of the population average leading height $\bar{y} = \frac{1}{n}\sum_{n} y$ = Phase II sample mean leading height b = least squares estimate of the regression relationship between x and y for the ground sampled polygons. $\bar{X} = \frac{1}{N}\sum_{N} x$ = Phase I population mean leading height $\bar{x} = \frac{1}{n}\sum_{n} x$ = Phase I sample mean n = number of ground sampled polygons N = number of polygons in the population where the Phase I age ≥ 20.
2. Compute the leading species (h) regression estimates $\hat{\bar{Y}}_h = \bar{y}_h + b_h(\bar{X}_h - \bar{x}_h)$	Estimation is by leading species group (LS) $\hat{\bar{Y}}_h$ = estimated population mean height for LS h. $\bar{X}_h = \frac{1}{N}\sum_{N} x_h$ = Phase I mean for LS h. $\bar{y}_h = \frac{1}{n}\sum_{n} y_h$ = ground sample mean height for LS h $\bar{x}_h = \frac{1}{n}\sum_{n} x_h$ = Phase I sample mean height for LS h Ground sampled polygons are weighted by their selection probability.
3. Compute r estimate by leading species $r_h = \frac{\hat{Y}_h}{X_h}$	$\hat{Y}_h$ = estimated population total height, for LS h $X_h = \sum_{N} x_h$ = Phase I total of LS h.
4. Compute the "k" adjustment factor (to force compatibility between the regression estimator for all species combined and the ratio of means estimators by LS) $k = \frac{\hat{Y}}{\sum_h \hat{Y}_h}$ $\hat{\bar{Y}}_h^* = k \cdot \hat{\bar{Y}}_h = k \cdot r_h \cdot \bar{X}_h$	$k = \frac{\hat{Y}}{\sum_h \hat{Y}_h} = \frac{area \cdot \hat{\bar{Y}}}{\sum_h area_h \hat{\bar{Y}}_h}$ $area_h$ = sum of areas of polygon with LS= h $\hat{Y} \equiv \sum_h \hat{Y}_h$ *
5. Adjust the LS height for each polygon $\hat{y}_{ih} = k \cdot r_h \cdot x_{ih}$	$\hat{y}_{ih}$ = estimate of the average height for polygon i with LS = h x_{ih} = Phase I leading sp0 height for polygon i with LS h.

Figure 4. A flowchart of the adjustment algorithm for leading species height.

1. Compute the population average ratio using a regression estimator $\hat{\bar{Y}} = \bar{y} + b(\bar{X} - \bar{x})$	$\hat{\bar{Y}}$ = estimate of the population average ratio between the second and leading species height $\bar{y} = \frac{1}{n}\sum_{n} y$ = Phase II sample mean b = least squares estimate of the regression relationship between x and y for the ground sampled polygons. $\bar{X} = \frac{1}{N}\sum_{N} x$ = Phase I population mean ratio $\bar{x} = \frac{1}{n}\sum_{n} x$ = Phase I sample mean n = number of ground sampled polygons N = number of polygons in the population where the Phase I age ≥ 20.
2. Compute the leading species (h) regression estimates $\hat{\bar{Y}}_h = \bar{y}_h + b_h(\bar{X}_h - \bar{x}_h)$	Estimation is by leading species group (LS) $\hat{\bar{Y}}_h$ = estimated population average ratio between the second and leading species height for LS h. $\bar{X}_h = \frac{1}{N}\sum_{N} x_h$ = Phase I mean ratio for LS h. $\bar{y}_h = \frac{1}{n}\sum_{n} y_h$ = ground sample mean height ratio for LS h $\bar{x}_h = \frac{1}{n}\sum_{n} x_h$ = Phase I sample mean height ratio for LS h Ground sampled polygons are weighted by their selection probability.
3. Compute r estimate by leading species $r_h = \frac{\hat{Y}_h}{X_h}$	$\hat{Y}_h$ = estimated population total height, for LS h $X_h = \sum_{N} x_h$ = Phase I total of LS h.
4. Compute the "k" adjustment factor (to force compatibility between the regression estimator for all species combined and the ratio of means estimators by LS) $k = \frac{\hat{Y}}{\sum_h \hat{Y}_h}$ $\hat{\bar{Y}}_h^* = k \cdot \hat{\bar{Y}}_h = k \cdot r_h \cdot \bar{X}_h$	$k = \frac{\hat{Y}}{\sum_h \hat{Y}_h} = \frac{area \cdot \hat{\bar{Y}}}{\sum_h area_h \hat{\bar{Y}}_h}$ $area_h$ = sum of areas of polygons with LS= h $\hat{Y} = \sum_h \hat{Y}_h$ *
5. Adjust the second species height ratio for each polygon $\hat{y}_{ih} = k \cdot r_h \cdot x_{ih}$	$\hat{y}_{ih}$ = estimate of the average height ratio for polygon i with LS = h. x_{ih} = Phase I leading sp0 height ratio for polygon i with LS h.
6. Adjust the second species height $\hat{hs}_{ih} = \hat{y}_{ih} \cdot ht_{ih}$	ht_{ih} = Phase I leading sp0 height for polygon i with LS h. $\hat{hs}_{ih}$ = Phase I second sp0 height for polygon i with LS h.

Figure 5. A flowchart of the adjustment algorithm for second species height.

4.3.4 Age

Age was adjusted in a manner similar to height. Only the leading species age was available so only the leading species age was adjusted.

4.3.5 Volume

Volume was adjusted by sp0. The population average was computed using a regression estimator. The volumes of the sp0s within polygons were adjusted using a ratio of means estimator by sp0. These ratio of means estimators were scaled to be compatible with the population mean. This was essentially the same adjustment algorithm as for basal area except that species composition was not affected.

4.3.6 Minor species

Two approaches were used to deal with minor species groups. For basal area and volume, the minor species groups were lumped into an "other" category and a single adjustment relationship estimated for the "other" category and used for all the species groups within it. Both volume and density by species have attribute estimates for all the species found in either Phase I or Phase II. In the Sunshine Coast District, 5 sp0 groups had more than 10 ground sampled polygons with positive Phase I estimates.

Age, height, and SI were only available for the leading species in Phase I. Some species with positive basal area were never the leading species and never had an estimate of age, height, or SI. Only 3 species groups had 10 or more ground sampled polygons with positive ages. These species groups with fewer observations were combined into an "other" species group.

4.3.7 Phase I species or the adjusted species

Should the original Phase I labels be used or the results of the sequential balancing used to determine the leading species and the subsequent basis for the age, height, and SI adjustment? The following discussion uses height as an example but the same arguments apply to age and SI.

The current photo interpretation procedures specify the height of the leading and secondary species are to be estimated. As an example, let the Phase I species composition of a polygon be Fd70Hw30 and following adjustment the species composition is Hw70Fd30. One either assumes the photo interpreter misidentified the species or estimated the composition incorrectly. The sequential balancing algorithm only adjusts the species composition. The problem of species misclassification is never directly addressed.

Assuming the photo interpreter was incorrect in the species composition, the height associated with Fd becomes the height of the secondary species. If the leading species changes, the height of the leading species also changes.

If the photo interpreter incorrectly identified Hw as FD, the leading species height remains the leading species height even if the leading species changes.

In about 7% of the area, the leading species changed following sequential balancing. The number of Phase II leading heights in the ground sampled polygons is low (96 observations in the Sunshine Coast District). All of these heights should be used to calibrate the adjustment relationship regardless of whether they remain the leading species following sequential balancing.

Recommendation – The default adjustment should use the adjusted species composition. To estimate the adjustment relationship, all the Phase I data should be used regardless of whether the species change order following adjustment. The Phase I height should remain with the species it was associated with before adjustment.

4.3.8 Missing data

If the species' order change following sequential balancing of basal area, the adjusted Phase I polygon labels may not include height or age of the leading or secondary species. This problem is unique to the ADJ_{sp0} algorithm. In the Sunshine Coast District, 6.9% of the area changed leading sp0 following sequential balancing. The height and age of the adjusted leading species are necessary. A number of methods are available for estimating the height, age, and site index of the adjusted leading species given that these data were available for the original leading species. The site index can be estimated from site

factors (e.g., Wang et al. 1994), the age assumed to be the same as the original leading species and the height estimated from the site index. Site index conversion equations are available for some species pairs (Nigh 1995a,b,c) and the site index for the new leading species estimated from the original leading species' site index. Both these methods introduce additional uncertainty into the Phase I estimates.

4.4 Species assignments

Generally, the Phase II data are collected by species, not genus. In Phase I, the data may be collected to the genus or species level. For instance, the species composition in Phase I is "H" rather than "HW" or "HM". It is recommended that adjustments are undertaken by species group. This will result in population estimates by species group. If the attributes are required by species, the species proportions in each species group can be estimated by the Phase II data.

If species composition is required at the species level for all polygons, species can be assigned by sorting the data (e.g., elevation, BEC, leading species, etc.) and genera assigned to species by ecological associations. These assignments could be implemented to match the Phase II proportions of the species within a genus. Development of rules for assigning species would require advice from experts in each inventory unit.

For illustrative purposes, species assignments were undertaken in the ADJ_{SQB} adjustment algorithm but, in the absence of expert rules, the assignments were arbitrary. The Phase I file was sorted by mapsheet and polygon and species were assigned in alphabetical order, i.e., "HM" assigned to the first records and "HW" to the last records with the "H" genus. Note that sorting by mapsheet and polygon added no new information to the data base. However, the algorithm would be improved considerably by more intelligent sorting. The current implementation (sorting by mapsheet and polygon) is not recommended but is included to illustrate how species could be assigned.

5. An example

The algorithms described in the previous section were applied to the Sunshine Coast District. The Phase I data were compiled to the 17.5cm utilization level. The 7.5cm utilization level was used for the Phase II data.

Of the 613,327ha (Table 1) in the vegetated treed stratum, 591,718ha had a Phase I age $\geq$ 20 years.

An artificial example is given in Appendix A.

5.1 ADJ_{LS}

5.1.1 Basal area

Basal area first was adjusted by leading species. The adjustment generally increased the basal area (Table 6). Note the fairly close mean basal areas for the Phase I population and the Phase I sample, except for the "Other" leading species. This gives some indication the sample is a good representation of the population. The overall population total is derived by adding the D, F, H, and Other strata rather than the "All" stratum. As is seen in Table 6, the relative prediction errors are considerably higher for the leading species strata than for the population as a whole. Although the ADJ_{LS} algorithm does not lead to substantially different population values, it does lead to larger prediction errors.

Table 6. The adjustment statistics are given by leading sp0 strata for basal area. Leading sp0s with less than 10 ground sampled polygons with a positive Phase I basal area were lumped into an "other" leading sp0. The F-statistic (F_{obs}) for the null hypothesis of a common regression line (vs. the alternative hypothesis of separate lines by sp0 Stratum) is 2.12. The $p(F_{6,94} > F_{obs}) = 0.0578$.

Leading sp0 stratum	Area (ha)	BA (m2/ha)			Number of polygons	Correlation	Intercept	Slope	Adjusted basal area (m2/ha)	Sampling error[4]
		Phase I population	Phase I sample	Phase II sample						
D	40725	32.38	33.53	40.64	12	0.57	17.26	0.70	39.83	27.8%
F	202940	37.19	38.76	52.95	37	0.59	20.78*	0.83**[5]	51.65	15.0%
H	212663	33.54	35.84	58.36	39	0.24	49.27**	0.25	57.77	18.1%
Other	135390	27.55	37.19	48.54	14	0.90	22.62**	0.70**	41.82	16.8%
All_{sum}[6]	591718								50.79	10.0%
All_{pop}[7]	591718	33.34	36.86	53.30	102	0.43	35.32**	0.49**	51.59	10.1%

5.1.2 Density

Density was adjusted in a manner similar to basal area. The adjustment results are given in Table 7. The decrease in D density was countered by an increase in all other densities. Note the relatively low correlations between the Phase I and Phase II densities particularly for all species combined leading to relatively large prediction errors. Once again, the strata are the basis for the population estimates in the ADJ_{LS} algorithms and the population prediction error is derived from the strata prediction errors. The adjustment relationships were dominated by the intercept term.

The relationship between the photo and ground density estimates is marginal and the merit of using the auxiliary (Phase I) is questionable. Using the ground sampling information alone leads to a marginally lower sampling error except for the "other" sp0 group. This is a result of the poor correlation between the Phase I and Phase II estimates and the relatively small sample sizes.

Table 7. The adjustment statistics are given by leading sp0 strata for density. Leading sp0s with less than 10 ground sampled polygons with a positive Phase I density were lumped into an "other" leading sp0. The F-statistic (F_{obs}) for the null hypothesis of a common regression line (vs. the alternative hypothesis of separate lines by sp0 Stratum) is 0.424. The $p(F_{6,94} > F_{obs}) = 0.861$.

Lead sp0 stratum	Area (ha)	Density (TPH)			Number of polygons	Correlation[8]	Intercept	Slope	Adjusted density (TPH)	Sampling error based on regression	Sampling error based on Phase II only
		Phase I popn	Phase I sample	Phase II sample							
D	40725	1082	951	935	12	-0.20	1439	-0.53	866	71.6%	64.5%
F	202940	862	975	1000	37	0.17	791**	0.21	975	27.5%	26.9%
H	212663	790	722	1110	39	0.02	1094**	0.02	1111	27.8%	27.5%
Other	135390	613	999	914	14	-0.13	1010**	-0.10	951	71.0%	71.6%
All_{sum}	591718								1011	18.1%	17.7%
All_{pop}	591718	794	875	1023	102	-0.02	1040**	-0.02	1025	17.6%	

[4] The sampling error is ½ the 95% confidence interval around the mean expressed as a percentage of the predicted mean. That is, $t_{n-2,0.025}$*SQRT(MSE/n)/Adjusted mean.

[5] A single asterisk indicates rejection of the null hypothesis (parameter = 0) at the 5% level of significance, a double asterisk indicates rejection at the 1% level of significance.

[6] The "All_{sum}" is based on the sum of the leading species strata.

[7] The "All_{pop}" is based on the all the polygons and was not used in the ADJ_{LS}. It included in the tables to enable comparison of the relative prediction errors.

[8] Note the low correlations. In this case, a better (lower variance) estimator would have been obtained just using the Phase II sample data.

5.1.3 Age

Age was adjusted in a manner similar to basal area. The adjustment results are given in Table 8. The correlations were low and the adjustment dominated by the intercept term. The prediction errors were relatively large.

Table 8. The adjustment statistics are given by leading sp0 strata for total age. Leading sp0s with less than 10 ground sampled polygons with a positive Phase I age were lumped into an "other" leading sp0. The F-statistic (F_{obs}) for the null hypothesis of a common regression line (vs. the alternative hypothesis of separate lines by sp0 Stratum) is 1.860. The $p(F_{6,94} > F_{obs}) = 0.096$.

Lead sp0 stratum	Area (ha)	Age (years)			Number of polygons	Correlation	Intercept	Slope	Adjusted age	Sampling error
		Phase I population	Phase I sample	Phase II sample						
D	40725	63	65	44	12	0.82	0.48	0.66**	42	19.4%
F	202940	127	100	108	37	0.73	34.43*	0.73**	128	14.8%
H	212663	201	195	184	39	0.65	30.18	0.80**	190	21.2%
Other	135390	227	204	137	14	0.75	44.80	0.45**	148	28.4%
All_{sum}	591718								149	11.8%
All_{pop}	591718	172	154	138	102	0.70	32.05*	0.69**	151	11.9%

5.1.4 Height

Height was adjusted in a manner similar to basal area. The adjustment results are given in Table 9. Similar to age, the overall average did not shift much but some species had significant shifts. The correlations were higher and the average adjustments were relatively minor by leading sp0 group.

Table 9. The adjustment statistics are given by leading sp0 strata for height. Leading sp0s with less than 10 ground sampled polygons with a positive Phase I height were lumped into an "other" leading sp0. The F-statistic (F_{obs}) for the null hypothesis of a common regression line (vs. the alternative hypothesis of separate lines by sp0 Stratum) is 2.500. The $p(F_{6,94} > F_{obs}) = 0.0274$.

Lead sp0 strata	Area (ha)	Height (m)			Number of polygons	Correlation	Intercept	Slope	Adjusted height	Sampling error
		Phase I population	Phase I sample	Phase II sample						
D	40725	23.8	24.9	27.1	12	0.48	8.29	0.76	26.3	24.0%
F	202940	29.1	26.3	28.3	37	0.77	4.83	0.89**	30.8	8.8%
H	212663	25.7	26.0	24.8	39	0.32	17.14**	0.29*	24.6	14.5%
Other	135390	24.6	25.4	24.5	14	0.83	7.30	0.68**	24.0	19.2%
All_{sum}	591718								26.7	7.0%
All_{pop}	591718	26.5	25.9	26.1	102	0.57	11.07**	0.58**	26.4	7.4%

5.1.5 Volume

Volume was adjusted in a manner similar to basal area. The adjustment results are given in Table 10. The adjustment caused reduced the average volume by leading sp0 group.

Table 10. The adjustment statistics are given by leading sp0 strata for volume. Leading sp0s with less than 10 ground sampled polygons with a positive Phase I height were lumped into an "other" leading sp0. The F-statistic (F_{obs}) for the null hypothesis of a common regression line (vs. the alternative hypothesis of separate lines by sp0 Stratum) is 3.137. The $p(F_{6,94} > F_{obs}) = 0.0076$.

Lead sp0 strata	Area (ha)	Volume/ha (m3/ha)			Number of polygons	Correlation	Intercept	Slope	Adjusted volume	Sampling error
		Phase I population	Phase I sample	Phase II sample						
D	40725	449.0	502.2	399.3	12	0.69	105.43	0.59*	368.2	31.2%
F	202940	557.7	502.8	511.7	37	0.78	63.74	0.89**	560.7	14.8%
H	212663	632.7	695.0	527.1	39	0.27	395.06**	0.19*	515.2	26.8%
Other	135390	538.2	645.7	459.8	14	0.82	152.86*	0.48**	408.6	35.2%
All_{sum}	591718								479.8	12.9%
All_{pop}	591718	572.7	606.4	499.1	102	0.50	271.71**	0.37**	486.4	13.9%

5.2 ADJ_{CONST}

5.2.1 Basal area

Basal area first was adjusted using all ground sampled polygons and then by leading species group. The adjustment generally increased the basal area (Table 11) particularly for F and H while others were reduced (e.g., D). The second last column is omitted in subsequent tables. The shifts in total basal area are given by leading species in Table 21. ADJ_{CONST} resulted in approximately 1.6% more adjusted basal area than ADJ_{LS} at the population level.

The constraint that the basal area by leading species be compatible with the computed population values increased the adjustment (see the last two columns in Table 11).

Table 11. The adjustment statistics are given by species strata for basal area. Leading species found in less than 10 of the ground sampled polygons were grouped into an "other" leading sp0

Lead Sp0 strata	Area (ha)	BA (m2/ha)			Number of polygons			Correlation	Ratio[9]	k[10]	Predicted BA		
		Phase I population	Phase I sample	Phase II sample	Phase I sample	Phase II sample	Phase I and Phase II				Unadjusted[12]	Adjusted[13]	Sampling error[11]
D	40725	32.38	33.53	40.64	12	12	12	0.57	1.21	1.08	39.24	42.19	30%
F	202940	37.19	38.76	52.95	37	37	37	0.59	1.37	1.08	50.81	54.63	15%
H	212663	33.54	35.84	58.36	39	39	39	0.24	1.63	1.08	54.61	58.71	22%
Other	135390	27.55	37.19	48.54	14	14	14	0.90	1.31	1.08	35.96	38.66	31%
ALL	591718	33.340	36.856	53.303	102	102	102	0.435			47.98[14]	51.59	10%

[9] The ratio is equal to phaseIIsamp / phaseIsamp. For species "D" in Table 11, the ratio is equal to (40.64/33.53) = 1.21.

[10] k = unadjusted total/adjusted total = 51.59/47.98 = 1.075.

[11] The mean squared error for the ratio estimate was estimated as $\sum \frac{(y-\hat{y})^2}{n-1}$ where $\hat{y} = r \cdot x$

[12] The unadjusted basal area is equal to the unadjusted ratio multiplied by the Phase I population basal area. . For species "D" in Table 11, the unadjusted basal area is equal to 32.38 * 1.21 = 39.24.

[13] The adjusted basal area is equal to the unadjusted basal area multiplied by the correction factor to ensure the species sum to the population total. For species "D" in Table 11, the adjusted basal area is equal to 39.24 * (51.59/47.98) = 42.19.

[14] Estimated as the sum of the unadjusted stratum means.

5.2.2 Density

Density was adjusted by leading sp0 group. The adjustment results are given in Table 12. The decrease in D and F density is countered by an increase in all other densities. Note the relatively low correlation between Phase I and Phase II density for all species combined. The adjustment led to considerably larger differences in the average density by species group than ADJ_{LS}. The negative correlations imply that as the Phase I estimates of TPH increase, the Phase II estimates decrease. This is contrary to what one would expect and may be a function of the estimation algorithm used. The correlations should improve (or at least become positive) if density is available directly from Phase I.

Table 12. The adjustment statistics are given by leading species strata strata for density. Leading species groups found in less than 10 of the ground sampled polygons were grouped into an "other" leading sp0.

Lead sp0 strata	Area (ha)	Density (TPH)			Number of polygons			Correlation	Ratio	k	Predicted TPH	Sampling error
		Phase I population	Phase I sample	Phase II sample	Phase I sample	Phase II sample	Phase I and Phase II					
D	40725	1082	951	935	12	12	12	-0.20	0.98	1.09	1159	64%
F	202940	862	975	1000	37	37	37	0.17	1.03	1.09	963	32%
H	212663	790	722	1110	39	39	39	0.02	1.54	1.09	1322	30%
Other	135390	613	999	914	14	14	14	-0.13	0.91	1.09	611	177%
All	591718	794	875	1023	102	102	102	-0.02			1025	18%

5.2.3 Age

Age was adjusted in a manner similar to density. The adjustment results are given in Table 13. Note the rather dramatic decrease in "Other" age (and to a lesser extent H) despite the relatively high correlation.

Table 13. The adjustment statistics are given by leading species strata for total age. Leading species groups found in less than 10 of the ground sampled polygons were grouped into an "other" leading sp0.

Lead sp0 strata	Area (ha)	Age (years)			Number of polygons			Correlation	Ratio	k	Adjusted Age	Sampling error
		Phase I population	Phase I sample	Phase II sample	Phase I sample	Phase II sample	Phase I and Phase II					
D	40725	63	65	44	12	12	12	0.82	0.67	0.98	41	18%
F	202940	127	100	108	37	37	37	0.73	1.08	0.98	135	11%
H	212663	201	195	176	39	39	39	0.62	0.90	0.98	178	16%
Other	135390	227	204	137	14	14	14	0.75	0.67	0.98	150	20%
All	591718	172	154	135	102	102	102	0.68	0.88		147	13%

5.2.4 Height

Height was adjusted in a manner similar to density. All polygons were used in each stratum. The adjustment results are given in Table 14. The overall average did not shift appreciably and the leading species group shifts were relatively minor.

Table 14. The adjustment statistics are given by leading sp0 strata for height. Leading species groups found in less than 10 of the ground sampled polygons were grouped into an "other" leading sp0.

Lead sp0 strata	Area (ha)	Height (m)			Number of polygons			Correlation	Ratio	k	Adjusted Height	Sampling Error
		Phase I population	Phase I sample	Phase II sample	Phase I sample	Phase II sample	Phase I and Phase II					
D	40725	23.8	24.9	27.1	12	12	12	0.48	1.09	0.99	25.6	24%
F	202940	29.1	26.3	28.3	37	37	37	0.77	1.07	0.99	30.8	8%
H	212663	25.7	26.0	23.6	39	39	39	0.38	0.91	0.99	23.0	14%
Other	135390	24.6	25.4	24.5	14	14	14	0.83	0.97	0.99	23.5	14%

Lead sp0 strata	Area (ha)	Height (m)			Number of polygons			Correlation	Ratio	k	Adjusted Height	Sampling Error
		Phase I population	Phase I sample	Phase II sample	Phase I sample	Phase II sample	Phase I and Phase II					
All	591718	26.5	25.9	25.6	102	102	102	0.58			26.0	8%

5.2.5 Volume

Volume was adjusted in a manner similar to basal area. The adjustment results are given in Table 15. Most volumes decreased as a result of the adjustment.

Table 15. The adjustment statistics are given by species strata for volume. Leading species groups found in less than 10 of the ground sampled polygons were grouped into an "other" leading sp0.

Lead sp0 strata	Area (ha)	Volume/ha (m3/ha)			Number of polygons			Correlation	Ratio	k	Adjusted volume	Sampling error
		Phase I population	Phase I sample	Phase II sample	Phase I sample	Phase II sample	Phase I and Phase II					
D	40725	473.7	529.7	399.3	12	12	12	0.69	0.75	1.01	357.1	362.1
F	202940	584.4	526.9	511.7	37	37	37	0.78	0.97	1.01	567.6	575.5
H	212663	600.7	661.5	527.1	39	39	39	0.28	0.80	1.01	478.6	485.3
Other	135390	476.0	564.7	459.8	14	14	14	0.82	0.81	1.01	387.5	392.9
All	591718	557.8	588.4	499.1	102	102	102	0.51			486.6	486.6

5.3 ADJ_{SQB}

The same data set was adjusted using the ADJ_{SQB} algorithm, i.e., first adjusting species composition using sequential balancing. Except for volume, the standard errors and confidence intervals for the computed population means are the same as in the ADJ_{CONST} algorithm.

5.3.1 Basal area

The basal area adjustment was carried out using sp0 as the grouping as described in 4.3. Note the overall computed BA in Table 16 is the same as in ADJ_{CONST} but the distribution by leading species has changed (Table 21). The overall matching success (proportion of polygons with a species present in both Phase I and Phase II) is considerably higher than for ADJ_{CONST}. It is more likely that a species is identified in a polygon in both Phase I and Phase II than a particular species is the leading species in both Phase I and Phase II. If a species was not present in a ground sampled polygon, it was assigned a basal area of zero. This affected the interpretation of the correlation coefficient so it is not presented here.

Table 16. The adjustment statistics are given by sp0 for basal area. Sp0s with less than 10 ground sampled polygons with a positive Phase I basal area were grouped into an "other" sp0.

Sp0 group	Sp0	Area (ha)	BA (m2/ha)			Number of polygons			Interc ept	Ratio or slope [15]	k	Adjusted BA	Sampling Error
			Phase I pop'n	Phase I sample	Phase II sample	Phase I sample	Phase II sample	Phase I and Phase II					
B	B	591718	3.43	5.02	6.96	102	102	102		1.39	1.06	5.02	39.7%
C	C	591718	4.31	3.84	8.58	102	102	102		2.23	1.06	10.16	22.1%
D	D	591718	2.10	2.38	1.72	102	102	102		0.72	1.06	1.60	77.5%
F	F	591718	10.43	10.37	12.84	102	102	102		1.24	1.06	13.63	17.9%
H	H	591718	10.56	12.01	18.19	102	102	102		1.52	1.06	16.89	14.1%
Oth	Oth	591718	0.56			102	102	102		2.34	1.06	1.37	74.7%
Oth	AC	591718	0.09	0.03	0.09	102	102	102		2.34	1.06	0.22	
Oth	AT	591718	0.00			102	102	102		2.34	1.06	0.00	
Oth	E	591718	0.01	0.04	0.10	102	102	102		2.34	1.06	0.02	

[15] The ratio estimator is given for the Sp0 groups and the regression slope is given for the population (All sp0 groups combined).

Sp0 group	Sp0	Area (ha)	BA (m2/ha)			Number of polygons			Interc ept	Ratio or slope $_{fs}$	k	Adjusted BA	Sampling Error
			Phase I pop'n	Phase I sample	Phase II sample	Phase I sample	Phase II sample	Phase I and Phase II					
Oth	MB	591718	0.27	0.27	0.48	102	102	102		2.34	1.06	0.66	
Oth	PA	591718	0.03			102	102	102		2.34	1.06	0.07	
Oth	PW	591718	0.03	0.10	0.24	102	102	102		2.34	1.06	0.07	
Oth	S	591718	0.14	0.10	0.35	102	102	102		2.34	1.06	0.34	
PL	PL	591718	0.47	0.83	1.06	102	102	102		1.27	1.06	0.63	
Y	Y	591718	1.48	1.85	2.70	102	102	102		1.46	1.06	2.29	47.9%
All	All	591718	33.34	36.86	53.30	102	102	102	35.32	0.49		51.588	10.1%

5.3.2 Density

The overall density adjustment used a least squares regression estimator. The polygon level density adjustment was carried out using the sp0 of the leading species (following basal area adjustment). The adjustment ratios are given in Table 17. The overall adjustment is the same as for ADJ_{CONST}. However, due to the sequential balancing, some of the leading species changed, leading to different adjustment relationships at the leading species level. Only 4 leading sp0s had 10 or more ground sampled polygons with a positive Phase I TPH. The basis for the adjustment has a much lower sample size and the correlations are much lower than for basal area.

Table 17. The adjustment statistics are given for trees per hectare (TPH) by leading sp0 (the sp0 of greatest basal area in the polygon). Leading sp0s with less than 10 ground sampled polygons with a positive Phase I TPH were lumped into an "other" leading sp0.

Lead sp0 group	Area (ha)	TPH (stems/ha)			Number of polygons			Intercept	Ratio or slope	k	Adjusted TPH	Sampling error
		Phase I pop'n	Phase I sample	Phase II sample	Phase I sample	Phase II sample	Phase I and Phase II					
C	591718	95	81	80	102	102	102		0.99	1.09	102	12.3%
D	591718	52	61	63	102	102	102		1.02	1.09	57	20.4%
F	591718	269	322	334	102	102	102		1.04	1.09	304	4.6%
H	591718	278	281	433	102	102	102		1.54	1.09	466	3.4%
Other	591718	100	130	113	102	102	102		0.87	1.09	95	14.5%
All	591718	794	875	1023	102	102	102	1039.86	-0.02		1025	17.6%

5.3.3 Age, height, and volume.

Ages and heights were adjusted by species group. As seen in Table 18 and Table 19, both were adjusted downwards. In future inventories based on new photography, the leading and secondary species will have age, height, and site index estimates. Having more data will allow for adjustment by sp0 or by leading sp0.

Volume was adjusted by sp0 (Table 20).

Table 18. The adjustment statistics are given for total age by leading sp0 (the sp0 of greatest basal area in the polygon). Leading sp0s with less than 10 ground sampled polygons with a positive Phase I and Phase II age were lumped into an "other" leading sp0.

Lead sp0 group	Area (ha)	Age (years)			Number of polygons			Intercept	Slope	k	Adjusted Age	Sampling error
		Phase I pop'n	Phase I sample	Phase II sample	Phase I sample	Phase II sample	Phase I and Phase II					
F	186857	127	97	125	36	32	22	48.29	0.76	0.83	121	6.8%
H	207882	202	192	166	37	37	16	-44.46	1.22	0.83	168	7.0%
Other	196979	184	167	100	29	35	12	-2.86	0.84	0.83	126	6.8%
All	591718	172	154	129	102	102	102	47.69	0.53		139	18.7%

Table 19. The adjustment statistics are given for height by leading sp0 (the sp0 of greatest basal area in the polygon). Leading sp0s with less than 10 ground sampled polygons with a positive Phase I and Phase II height were lumped into an "other" leading sp0.

Lead sp0 group	Area (ha)	Height (m)			Number of polygons			Intercept	Slope	k	Adjusted Height	Sampling error
		Phase I pop'n	Phase I sample	Phase II sample	Phase I sample	Phase II sample	Phase I and Phase II					
F	186857	28.9	31.7	26.0	36	32	22	9.03	0.9	0.88	30.8	9.0%
H	207882	25.5	24.9	25.4	37	37	16	-2.69	0.97	0.88	19.5	18.1%
Other	196979	25.2	26.5	19.7	29	35	12	-2.35	1.19	0.88	24.2	35.4%
All	591718	26.5	25.9	24.4	102	102	102	12.28	0.47		24.6	11.5%

Table 20. The adjustment statistics are given for volume by sp0. Sp0s with less than 10 ground sampled polygons with a positive Phase I volume were lumped into an "other" sp0.

Lead sp0 group	Area (ha)	Volume (m3/ha)			Number of polygons			Intercept	Ratio or slope	k	Adjusted volume	Sampling error
		Phase I pop'n	Phase I sample	Phase II sample	Phase I sample	Phase II sample	Phase I and Phase II					
B	591718	65.5	81.8	65.0	102	102	102		0.79	1.00	52.2	11.6%
C	591718	86.6	74.9	69.7	102	102	102		0.93	1.00	80.8	8.9%
D	591718	16.7	19.2	20.5	102	102	102		0.90	1.00	15.2	28.0%
F	591718	151.7	133.0	135.2	102	102	102		1.01	1.00	153.5	5.3%
H	591718	155.4	170.2	171.9	102	102	102		1.01	1.00	157.5	5.2%
Other	591718	15.8	15.0	11.6	102	102	102		0.77	1.00	12.2	19.6%
PL	591718	5.7	10.0	6.3	102	102	102		0.63	1.00	3.6	65.6%
Y	591718	21.8	27.2	23.4	102	102	102		0.86	1.00	18.8	19.3%
All	591718	519.2	531.2	499.1	102	102	102	265.45	0.44		493.8	13.4%

Note the Phase I estimates of volume (from VRIStart based on the adjusted basal area, density, age, and height) in ADJ_{SQB} and ADJ_{CONST} are considerably higher than in ADJ_{LS}.

6. Discussion

6.1 Defining the population

Defining the population to be adjusted is not a trivial task. The adjustment algorithms presented here focus on the tree attributes and it was logical to restrict the population to the vegetated treed portion of the data set. The ground sampling focuses on trees >4cm dbh while the Phase I inventory theoretically includes all trees >0cm dbh. This brings up a number of issues. Classification of polygons as vegetated treed or not vegetated treed based on the photos or based on the ground sampling is not always the same. The ground information is generally considered more accurate. It is possible to correct the Phase I stratum means and totals using the Phase II information (see Schreuder et al 1993 section 5.2.2). However, it is more difficult to reconcile the Phase I polygon information to the corrected stratum estimates. For instance, if the ground sampling indicates a lower proportion of area in the vegetated treed stratum, it is possible to revise the stratum means. However, it is not obvious which polygons with a Phase I designation of vegetated treed should be shifted out of the stratum.

Similarly, the 4cm dbh lower limit for the ground sampling implies sort the of adjustment relationship given in Figure 6.

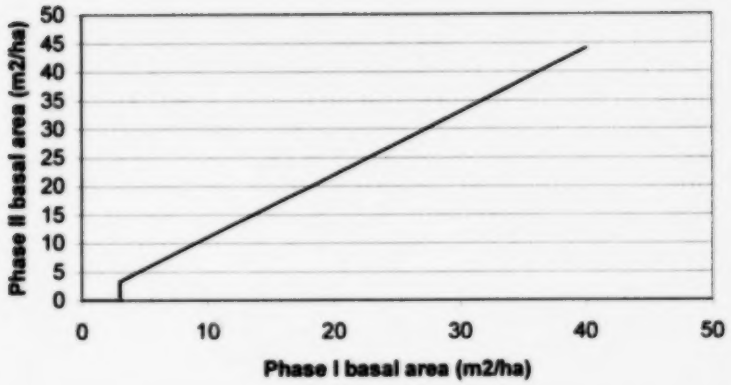

Figure 6. A theoretical example of the effect of the 4cm dbh lower limit in the ground sampling.

This sort of discontinuity in the adjustment relationship is acceptable if the attribute of interest is the basal area (and volume) of trees greater than 4cm dbh. However, if the attribute of interest is the basal area of all trees with a dbh >0cm, the adjustment relationship in Figure 6 is not adequate.

Several options are available. Option 1 is to remove the data points with the lower basal areas that have a higher incidence of small trees. By only including the polygons with higher Phase I basal areas, it is likely the Phase I basal area is dominated by large trees. However, this option assumes the relationship estimated from the higher basal areas can be extrapolated to lower basal areas.

Option 2 is to add some sort of factor to the Phase II estimates that corrects for the exclusion of trees <4cm dbh so the revised estimate of the Phase II basal area can be considered to include all trees.

Option 3 is to compile the volumes to a higher utilization limit. For instance, for the Sunshine Coast District, the primary utilization level was 17.5cm. Both the Phase I and Phase II data could be compiled to this utilization limit and the other utilization limits estimated as a function of the primary utilization level. This does not really alleviate the discontinuity problem in the adjustment relationship, it just shifts it along the x-axis in Figure 6.

Option 4 is to redefine the population to exclude the young stands with small trees. In this report, the population was defined as the polygons in the vegetated treed stratum that had a Phase I age of 20 years or more. This removed many of the polygons with no basal area. The disadvantage of this option is the stratification is based on an attribute measured with error (the Phase I age). The advantage is it removes many of the difficulties associated with the young stands which include the discontinuity in the adjustment relationship (Figure 6) and the high spatial variability associated with young stands. The high variability in young stands make them difficult to photo interpret and field sample with limited resources. The cutoff of 20 years used in this report is somewhat arbitrary but satisfies the objective of removing most of the polygons with small tree sizes. The cutoff could be increased or decreased if additional considerations became important.

6.2 Comparison between algorithms

All the algorithms lead to significant increases in basal area and in volume (see Table 21 and Figure 7). Assuming the Phase II sampling yields more accurate estimates of basal area and volume, these differences indicate an unadjusted Phase I inventory significantly underestimates basal area and volume. The different adjustment relationships by leading species or species group indicate that a single population correction factor may be inadequate. The density and age adjustments were also significant. The height adjustments were relatively minor.

Table 21. The effect of the various adjustment algorithms on basal area are given by leading species. Note that The ADJ_{SQB} algorithm can change the species composition and thus the leading sp0.

Leading Sp0	Area (ha) (Phase I)	Original BA (m^2)	ADJ_{LS} BA (m^2)	ADJ_{CONST} BA (m^2)	ADJ_{SQB} BA(m^2)	Area (ha) (ADJ_{SQB})
AC	2,540	47,676	72,376	66,908	152,110	3,737
AT	15	248	377	348	470	15
B	74,562	1,529,938	2,322,570	2,147,106	2,203,023	72,735
C	33,025	1,319,435	2,003,009	1,851,687	4,764,867	61,350
D	40,725	1,318,479	1,622,275	1,718,241	886,937	29,164
E	106	2,119	3,217	2,974	12,584	278
F	202,940	7,547,226	10,481,567	11,086,474	9,778,774	186,857
H	212,663	7,132,653	12,286,326	12,486,223	11,071,836	207,882
MB	2,824	98,796	149,981	138,650	414,768	7,026
PA	549	0	0	0	20,056	775
PL	8,316	158,584	240,743	222,555	216,532	8,412
PW	48	365	554	512	2,391	142
S	958	38,223	58,025	53,641	127,053	1,504
Y	12,447	534,383	811,237	749,950	873,874	11,841
ALL	591,718	19,728,126	30,052,256	30,525,273	30,525,273	591,718

The ADJ_{LS} and ADJ_{CONST} algorithms led to surprisingly similar results by leading species strata (Table 21 and Figure 7). The ADJ_{SQB} algorithm brought the species composition considerably closer to the Phase II species composition particularly for C and F.

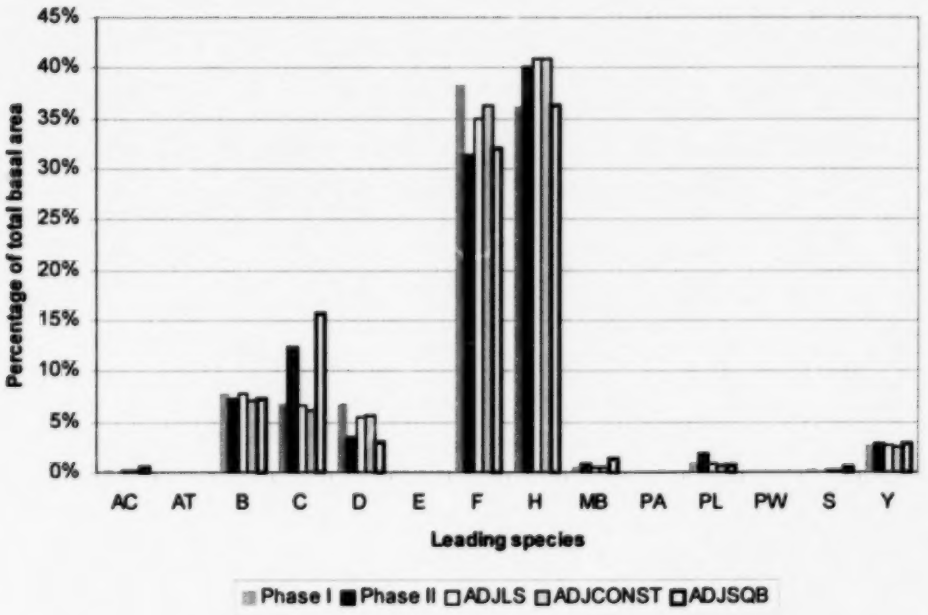

Figure 7. The basal areas in Table 21 are expressed as percentages of the population.

The most dramatic impact of all algorithms was the general increase in basal area and volume. The shifts in basal area by leading species or species group were relatively minor compared to the overall increases. This indicates that although there are some Phase I biases by species, the overall biases (all species combined) dominate.

Table 22. The effect of the various adjustment algorithms on basal area are given by sp0.

Sp0	Area (ha) (Phase I)	Original BA (m^2)	ADJ_{LS} BA (m^2)	ADJ_{CONST} BA (m^2)	ADJ_{SQB} BA(m^2)
AC	7,299	51,882	75,063	72,567	128,194
AT	15	134	203	188	330
B	235,169	2,030,671	3,279,849	3,194,779	2,971,337
CW	308,748	2,551,070	3,897,246	3,880,674	6,013,361
DR	116,184	1,243,603	1,628,962	1,707,993	947,660
E	1,302	5,518	7,848	7,854	13,635
FD	353,173	6,173,265	8,824,938	9,233,164	8,063,217
H	506,006	6,246,983	10,099,404	10,220,919	9,992,591
MB	26,113	156,923	223,876	222,764	387,742
PA	10,368	17,193	26,537	25,042	42,481
PL	48,816	277,201	407,063	405,870	370,278
PW	5,592	16,720	25,433	25,986	41,314
S	13,938	80,875	123,978	121,609	199,835
YC	138,982	876,087	1,431,854	1,405,853	1,353,298
All	1,771,705	19,728,126	30,052,256	30,525,263	30,525,273

6.3 Properties of the estimates

In Tables 6 through 20, the sampling errors associated with strata or species groups are higher than those associated with the population estimates. This is largely due to higher variances associated with the strata and species group estimates compared to the population estimates. Variances associated with individual polygon estimates are considerably higher than the strata and species group estimates. Therefore, if a single polygon is audited, the difference between the audit estimates and the inventory estimates may be quite large. This should not be considered evidence of inaccurate or imprecise population estimates. As the number of samples in the audit increases, the average difference between the audit and the inventory should decrease. The adjustment provides good population and stratum or species group estimates and attempts to provide reasonable polygon estimates. This provides a reasonable basis for strategic level decision-making and modeling.

6.3.1 Appropriate estimators

As discussed in section 3.4, one estimator will not be best for all attributes all of the time. For instance, for density (Table 7), the correlations are low and the sample sizes small. The sampling error associated with the regression estimator is generally larger than the inflation estimator. Thus, it appears the inflation estimator may be more appropriate for density than the regression estimator. In general, the regression estimator was more appropriate than the ratio estimator for the Sunshine Coast District dataset (except for density as noted previously) as seen in Table 6 and Tables 8 –10 where the intercept is generally positive and often statistically significant. The variance of the dependent variable around the regression line is also important when selecting from the family of regression estimators. For the ADJ_{SQB} algorithm, there is some indication of increasing variance with the independent variable for age and possibly height (Figure 8). However, the trends for basal area, density and volume are not clear.

Based on the Sunshine Coast District dataset, it appears that a regression estimator is often appropriate but the regression statistics should be examined to determine the significance of the regression and the individual parameters and the data should be plotted to examine the variance trends. For density, the regressions were not significant and the inflation estimators more efficient. For age and height, it appears the increase in Phase II variance with increase in Phase I may need explicit consideration in the adjustment model.

6.3.2 Bias of Phase I estimates

The bias associated with the Phase I estimates can be examined for the Sunshine Coast District dataset. If there is no bias in the Phase I or Phase II estimates, the relationship between Phase I and Phase II will have an intercept of zero and a slope of one. Assuming the Phase II estimates are not biased, this can be tested. For the ADJ_{LS} algorithm, the significance of the intercepts (i.e., whether the intercept is significantly

different from zero) is indicated in Tables 6 through 10. Generally, the intercepts are positive and in over half the cases significant at the 5% level. The null hypothesis of the slope being equal to 1 was tested for the attributes in tables 6 through 10. In 17 out of 25 adjustment relationships, the slope was significantly different from one. This indicates significant bias in the Phase I estimates.

6.3.3 Population and stratum estimates

All the adjustment procedures provide unbiased and minimum variance estimates at some level provided some assumptions are met. With ADJ_{LS}, the estimates are unbiased at the stratum level (leading sp0 stratum) while ADJ_{CONST} and ADJ_{SQR} are unbiased at the population level. The population estimate for ADJ_{LS} is the sum of unbiased, independent estimates and is therefore unbiased. This raises the question of the value of imposing the population constraint. Given the strata are relatively independent, the expected value of the difference between the adjusted population and the sum of the strata adjustments is approximately zero. In addition, the sampling errors associated with the sum of the strata and the population estimates are roughly equivalent (All_{sum} vs. All_{pop} in Tables 6 through 10). The population constraints add marginally to the complexity of the adjustment but do not affect the statistical properties. Thus the decision to impose the constraint can be based on external (non-statistical) considerations.

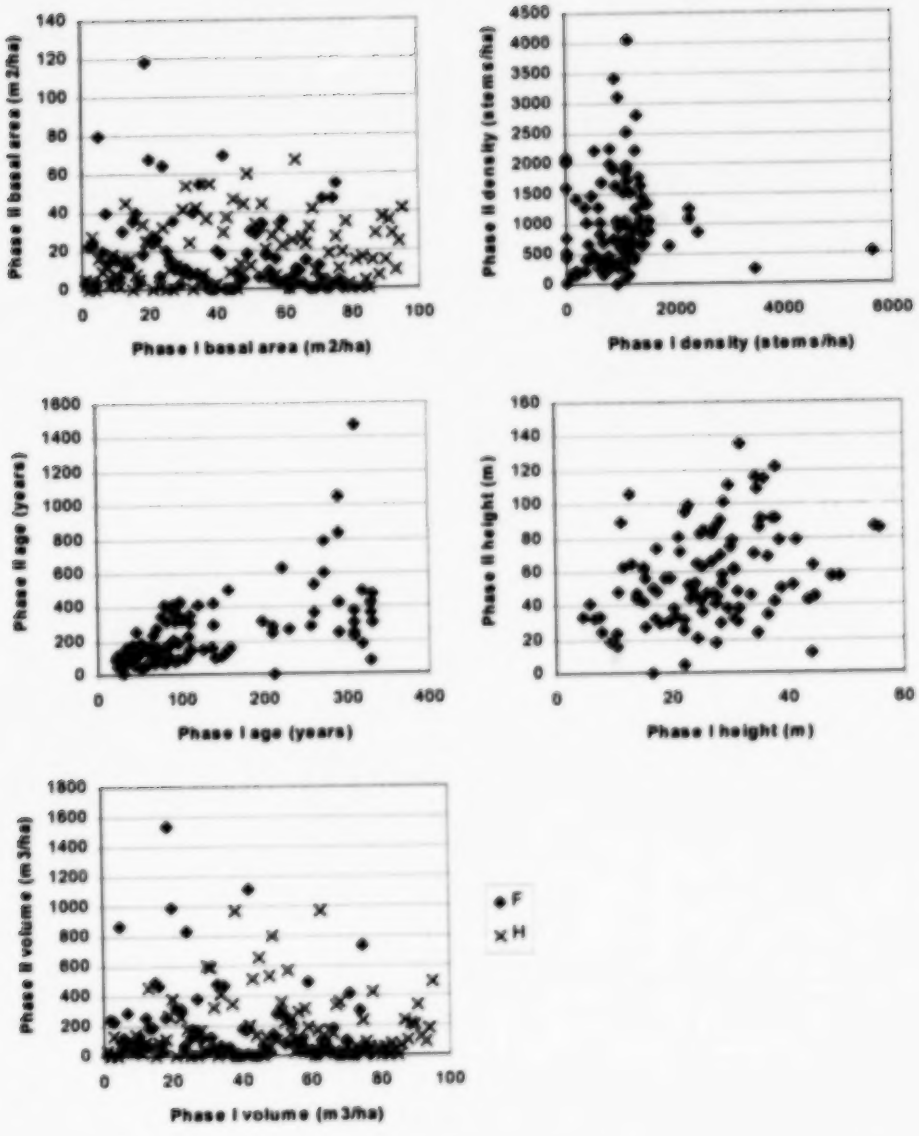

Figure 8. The relationship between Phase I and Phase II is shown for sum attributes in the ADJ_{SOB} algorithm.

ADJ_{LS} and ADJ_{CONST} do not address the issue of species misclassification so the population estimates should not be considered unbiased estimates of the species attributes. The attribute estimates are unbiased by strata (leading species group). ADJ_{SOB} does estimate the species groups separately so the population attribute estimates are unbiased by species group. The application of the correction factor k from equation (7) does affect the unbiased nature of the species population estimates but ensures an unbiased total for all species combined.

It should be emphasized that the adjusted estimates are unbiased estimates of the Phase II attributes. If, for example, the Phase II measurements of basal area were biased, the resulting adjusted estimates are unbiased estimates of the Phase II basal area but biased estimates of the true basal area.

6.3.4 Polygon estimates

The properties of the polygon estimates are not clear. Ratios were used to adjust attributes at the polygon level. This avoided the complications arising from positive and negative intercepts. Positive intercepts introduce species to all polygons in the stratum regardless of whether they were identified in the polygon in Phase I. Negative intercepts can produce negative attribute estimates that are not plausible.

In most cases, the intercept estimates were positive (Table 6 to Table 10). Both the ratio and the regression adjustment go through the means of the ground sampled polygons (see Figure 9). The regression also goes through a generally positive intercept while the ratio estimator goes through the origin. This leads to a relatively greater range in predictions at the polygon level using the ratio estimator rather than the regression estimator. In Figure 9, the regression estimates range from 17 to 52 m^2/ha in the range Phase I basal areas of 0 to 50 m^2/ha. The ratio estimates range from 0 to 60 m^2/ha.

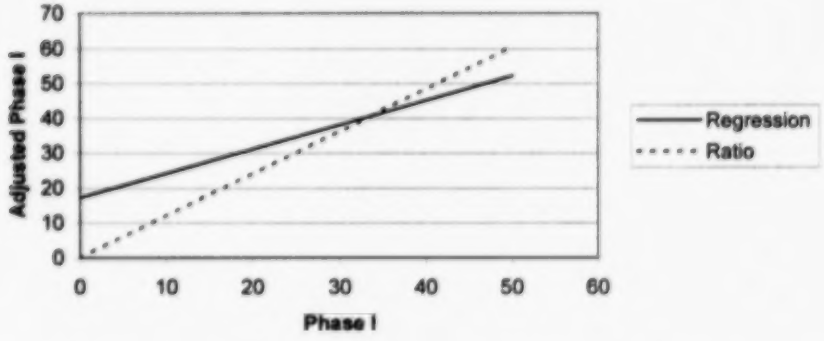

Figure 9. The regression estimator for stratum D from Table 6 is compared to a ratio estimator.

The greater range associated with the ratio estimators make it more likely that the predictions will exceed the range of the data, resulting in implausible predictions. Thus, although the population estimates have excellent statistical properties and the polygon estimates are compatible with the population estimates, the polygon estimates may not always be plausible. This situation is probably aggravated in the ADJ_{SOB} algorithm where the species' basal areas and volumes are adjusted independently.

The properties of the adjusted polygon estimates are important because they are used in the derivation of the adjusted polygon volumes. It is not necessary that the unadjusted Phase I volumes (based on VRIStart) be unbiased. It is important that they are correlated, hopefully well-correlated, with the Phase II estimates of volume.

6.4 Species misclassification

One of the more difficult issues is that of species misclassification, where one species is confused and interpreted to be one or more other species. No procedure has been developed to track such confusion in a meaningful way and switch species where it is necessary to do so. The algorithms implemented here do not address errors in the estimation of species composition in Phase I. The species identified within a polygon in Phase I are maintained, none are introduced or eliminated. In ADJ_{LS} and ADJ_{CONST}, the Phase I species composition is maintained. ADJ_{SOB} alters the species composition indirectly, by altering the relative amounts of the species identified in Phase I within a polygon. The basal area of each species group is adjusted independently and the species composition is computed from the results. A direct method of addressing species misclassification would involve reassigning a portion of one species' basal area within a polygon to another species within the same polygon. This is an extremely difficult problem. Solutions to

reduce the problem will likely involve using expert knowledge of the silvics of the species and of photo interpretation errors. The solutions will likely need to be customized for each inventory unit and generally involve rather complex algorithm to maintain compatibility with the population estimates and compatibility between variables. It is unlikely that purely statistical solutions will be adequate.

At some point, the issue of species misclassification should be addressed. It will not affect the adjusted population estimated but will affect the adjusted polygon estimates. Since these adjusted polygon estimates are the basis for any inventory updating or projection, the issue of species misclassification will affect updating and projection.

6.5 Species labels

The Phase I species composition sometimes includes genera designations rather than species, particularly for species within the same genus (HM and HW; BA, BG, and BL; SE, SS, and SW). The algorithm used in ADJ_{SQB} to allocate genera basal area to species can be used with any of the adjustment algorithms. Expert rules based on local knowledge of species association should provide better results.

6.6 Compatibility among various utilization limits

As Jim Flewelling has indicated (memo to Sam Otukol dated Feb. 23, 2000) it is fairly simple to create compatible adjustments for any component to the various utilization levels. His recommendation of using a logit function will be adopted. Volume – net decay waste and breakage (NDWB) is the only category 1 variable with multiple utilization levels.

6.7 Compatibility between attributes

Basal area (BA), density (TPH), and quadratic mean diameter (DQ) are related to one another by definition, i.e., $BA = TPH \times DQ^2 \times 0.00007854$. This relationship is nonlinear in DQ. Consider a population of two stands of equal area, one with a BA of 40 m^2/ha and TPH of 1000 stems/ha (and a DQ of 22.57 cm) and the other with a BA of 80 m^2/ha and 1000 stems/ha (and a DQ of 31.92 cm). The population average BA is 60 m^2/ha and the average TPH is 1000 stem/ha. The population DQ is 27.64 cm (based on the BA and TPH). However, the arithmetic average DQ is 27.24. For attributes that are nonlinearly related, one cannot expect compatibility at the polygon level, at the population level, and between the polygons and the population level for all the attributes.

Age, height, and site index also have a non-linear relationship.

6.8 Young stands

The adjustment focuses on polygons with appreciable volume and basal area. For recently disturbed stands without appreciable volume and basal area, the adjustment is based on little data (Jim Flewelling has not included polygons with an adjusted top height of less than 6m due to a shortage of data). Neither basal area nor volume may be a good basis for species composition for these stands (possibly stocking or density would be better). This is not addressed in this report. These stands are highly variable spatially and temporally and difficult to photo interpret. Data for these polygons may be available from reconnaissance surveys (e.g., following fire), regeneration surveys, etc.

The phase I information for young stands generally comes from silviculture survey data and has relatively accurate height and age data. The update algorithms may introduce considerable noise.

6.9 Incorporation of the adjustment into the database

The incorporation of the adjustment into the database is a relatively straightforward exercise. The adjustment ratios and correction factors (k) for attribute and strata can be maintained in lookup tables. The Phase I estimates for basal area, age, height, and density can then be linked to these tables and the adjusted values derived. These adjusted attributes are then input into VDYP7 and the resulting estimates adjusted by the volume lookup table of adjustment ratios and correction factors.

If the adjusted values for some attributes exceed acceptable bounds (see discussion in section 4.1.7), an

additional field should be maintained along with the Phase I files for those attributes. The field would contain an additional multiplier for those polygons exceeding bounds. Assuming any post-adjustment increase or decrease in one polygon would be balanced by a decrease or increase in one other polygon, the number of records requiring the additional multiplier should be small.

7. Outstanding issues

Some outstanding issues remain. Generally, they deal with minor species or a small number of polygons. The ground sampling is designed to produce unbiased population estimates. As a consequence of the relatively small number of ground sampled polygons and the diversity of the population, the ground sampled data do not provide an adequate base for dealing with minor species or relatively rare events.

7.1 Polygons with no basal area

In general, the Phase I estimates included proportionally more polygons with zero basal area than the Phase II sampling. Of the 102 ground-sampled polygons, 14 had zero basal area in Phase I and one had zero basal area in Phase II including one polygon with zero basal area in both phases. VDYP7 does compute volumes of veteran layers and these should be calculated and included in the adjustment process. The average basal area for the ground-sampled polygons with positive basal area was closer ($42.7m^2$/ha in Phase I vs. $53.8m^2$/ha in Phase II) than the average basal area for all polygons. Thus it seems that some basal area may be reasonably allocated to polygons with a Phase I estimate of zero basal area.

Two options are available.

- Option 1. Treat the zero basal area polygons as part of population and use a regression estimator for the adjustment. In general, the regression estimators had a positive slope that would introduce basal area to polygons where the Phase I basal area was estimated as zero. The regression estimators would have to be constrained so the intercept was non-negative. The species composition of the polygon would be taken from the Phase I data.
- Option 2. Treat the zero basal area polygons as a separate stratum. This may be justified on the basis of the wide variation in the polygons with zero basal area. They range from recently disturbed to low productivity to severely understocked. Each of these conditions may result in a different basal area. Rather than estimating the basal area from the regressions in Option 1, a separate average basal area could be estimated for each category of polygon with zero basal area. It is possible that some categories would have an estimate of zero basal area. The species composition could also be estimated from the ground sampling.

The feasibility of Option 2 may be limited by the availability of data. Note that introducing basal area to all polygons may not be acceptable since there may be polygons in the vegetated treed stratum which do not have any basal area.

7.2 Less common species

For various reasons, the inventory is best at identifying the presence of common, dominant species. These species are visible on the aerial photographs and are generally included in the ground sampling. Less common species may be overlooked on the photographs and may not occur in the ground sample polygons. There is no mechanism for dealing with species encountered in the field (e.g., en route to a ground sample cluster) but not tallied on the plot. Some species may only be identified in ground sampling. For instance, in Boston Bar yew was not identified in Phase I and was only found in one ground sample polygon. The amount of yew in the population may be estimated from the ground sample but the task of allocating the yew basal area to specific polygons is very difficult. One option is to distribute the yew basal area to all the polygons. This is unlikely to be an acceptable solution since the rarest species would occur in all polygons and the more common species in only a subset of the polygons. As well, carrying minor amounts of species in each polygon will increase the size of the database and affect the processing time of any updating algorithms. Alternatively, one could rank all the polygons in the population in terms of their "closeness" (e.g., in leading species, age, height, stocking, elevation, etc.) to the ground sampled polygon containing yew. The highest ranked polygon would be the most likely to contain yew. However, a single polygon does not provide a good basis for allocating species. Expert opinions or focused studies should provide

better information on species occurrence.

Regardless of the option implemented, users of the inventory need to be cautioned that the accuracy of the spatial distribution of minor species in the inventory is severely limited by the Phase II sample design and size.

7.3 Plausible estimates

The determination of plausible limits from the data for the adjusted attributes is difficult. The basic premise of the adjustment is that Phase I is biased and requires adjustment so the Phase I attributes alone cannot be used to determine acceptable attribute bounds. The Phase II sample size is relatively small and may not provide a good basis for estimating characteristics such as the minimum or maximum value of an attribute.

Routine data checking of the Phase I data should include some bound checks. These could also be used for the adjusted values. If data screening does not include bound checking, unreasonable adjusted attributes could be the result of extrapolating the adjustment relationships or unreasonable input data. In this case, all questionable adjusted values should be examined. If the Phase I data is suspect, that should be addressed. If the adjustment algorithm causes plausible Phase I values to become implausible, the adjusted value could be reset to the bound. Regardless, the suspect polygons should be examined individually rather than simply resetting the values to the bounds.

The problem of exceeding plausible limits becomes even more complex when one considers correlated attributes such as height and age. A height of 30m is plausible, as is an age of 10 years. But the combination is highly improbable. Ideally, a data checking algorithm would identify improbable combinations of attributes.

7.4 Customization or localization of the adjustment

The adjustment is very general and is designed to address the biggest problems, most of the time. As a result, the small, less frequent problems may be inadequately addressed. In some situations, it may be desirable to modify the adjustment algorithm. For instance, dominant species may be best adjusted using a ratio of means estimator. If the species is not observed in Phase I, it likely is not present so a zero intercept is reasonable. For minor, shade tolerant species, a regression estimator may be more appropriate. Some of the fir species may be frequently present even when not observed in Phase I. A positive intercept seems reasonable in these cases. Perhaps a non-linear relationship will be suggested for some data. It is beyond the scope of this report to recommend specific estimators for species or strata. This report has demonstrated that various estimators can be used to adjust Phase I. The algorithms presented here are general and should provide good population estimates. The adjusted polygon estimates are reasonable attempts to allocate the adjusted population values to the polygons, in the absence of any information but the Phase I and Phase II data. That is not the same as saying the adjusted polygon estimates are "best". If some polygons are more accurately inventoried as part of a separate study, those estimates should replace the Phase I adjusted attributes. If expert opinion indicates that all Douglas-fir in Boston Bar is coastal Douglas-fir, that should be incorporated.

This report identifies several general algorithms for adjusting the Phase I estimates. They represent a compromise between a general, robust solution and a more specific, local solution.

8. Conclusions

The adjustment of the possibly biased Phase I census to the unbiased Phase II sample has the potential to provide unbiased population estimates. The adjustments are significant, indicating the Phase I estimates alone may not be a good basis for decision-making. All attributes (with the possibly exception of height) were significantly affected by the adjustment.

The adjustment algorithm should include stratification by leading species or a separate adjustment by species group. The basis for the stratification must come from Phase I.

The Phase II sampling concentrates on trees >4cm dbh. This necessitates careful definition of the

population of interest. In this exercise, the adjustment was restricted to polygons in the vegetated treed stratum with Phase I ages of at least 20 years.

The adjustments by leading species group (ADJ_{LS} and ADJ_{CONST}) provide good results by leading species. They do not provide unbiased population estimates by species group. The more complex sequential balancing algorithm (ADJ_{SQB}) does provide unbiased population estimates by species group.

The adjustment is complicated by the additional objective of obtaining relatively good polygon estimates, some by species. The sample size for estimating a species-level adjustment relationship decreases as a result of differences in Phase I and Phase II species composition. In order to obtain good polygon estimates, the issue of species misclassification will need to be addressed. A statistical solution is unlikely to provide significant improvements. Expert opinion will likely be required to provide custom solution for each inventory unit.

The population estimates are the most precise. The stratum and species group estimates are less precise. The individual polygon estimates are the least precise. The difference between an audit and the inventory for a single polygon may be large but the difference between an audit of the population and the inventory population estimate should be small.

The consistency requirements between the variables are not strict but better expressed in terms of plausibility. A particular combination of age and height for a given species may have an average SI as predicted by SiteTools©. The closer the adjusted SI is to the SiteTools© prediction, the more plausible the adjusted value becomes. Extreme departures from the SiteTools© prediction may be flagged as extremely unlikely. These extreme values may be modified partly because they are unlikely but also because they may cause erratic predictions when input into VDYP7.

9. Literature Cited

Andersson, C. 1994. On the use of two-phase sampling in estimation of parameters in domains where data contain misclassification and measurement errors. Cta Univ. Ups. Studia Statistica Upsaliensia 3.. Uppsala, Sweden. 120pp

B.C. Ministry of Forests. 1997. Incorporation of the inventory audit in the Fraser timber supply area timber supply review. B.C. Ministry of Forests report, May 1997.

B.C. Ministry of Forests. 1999a. Vegetation Resources Inventory: Photo interpretation procedures V. 2.2. Prepared by the Ministry of Forests, Resources Inventory Branch for the Terrestrial Ecosystem task Force Vegetation Resources Inventory Committee, April 28, 1999. http://www.for.gov.bc.ca/ric

B.C. Ministry of Forests. 1999b. Report of the VRI Attributes Adjustment Working Group. December 14, 1999.

Cochran, W.G. 1977. *Sampling Techniques*, third ed. John Wiley & Sons. New York. 428p.

Iles, K. 1996. Comments on adjustment of volumes, by species, in polygons of the new Vegetation Inventory. Paper prepared for Ministry of Forests, Resources Inventory Branch.

J. S. Thrower and Associates Ltd. 1998. Vegetation Resources Inventory Statistical Analysis, Final Report. Prepared for the B.C. Ministry of Forests, Resources Inventory Branch, Project MFI-401-038.

Nigh, G.F. 1995a. Site index conversion equations for mixed Sitka spruce/Western hemlock stands. B.C. Min. For., Res. Br., Victoria, B.C., Ext. Note 02.

Nigh, G.D. 1995b. Site index conversion equations for mixed species stands. B.C. Min. For., Res. Br., Victoria, B.C., Res. Rep. 01. 20p.

Nigh, G.D. 1995c. The geometric mean regression line: a method for developing site index conversion equations for species in mixed stands. For. Sci. 41(1):84-98.

Penner, M. 1996. Optional adjustment procedures for the British Columbia Vegetation Resources Inventory. Prepared for the B.C. Ministry of Forests. 12p.

Penner, M. 2000. Species composition adjustment. Prepared for the B.C. Ministry of Forests, Resources Inventory Branch.

Schreuder, H.T., T.G. Gregoire, and G.B. Wood. 1993. *Sampling Methods for Multiresource Forest Inventory*. John Wiley & Sons, Inc. New York. 446p.

Wang, Q., G.G. Wang, K.D. Coates, and K Klinka. 1995. Use of site factors to predict lodgepole pine and interior spruce site index in the sub-boreal spruce zone. B.C. Min. For., Res. Note No. 114.

10. Apppendix A: An example of the various adjustment algorithms.

10.1 The data

The artificial data set was provided by Jim Flewelling. The raw (unadjusted) data are given in Tables 1-3. The leading species in a polygon was determined as the species with the most basal area. In the case of a tie, one species was selected arbitrarily. In practice, the Phase I leading species is identified in Phase I. If there is a tie for basal area in Phase II, the species with the greatest volume is selected.

Table 23. The Phase I ground sampled polygons.

Poly #	Polygon area (ha)	Leading species	Basal area by species (m^2/ha)				Polygon basal area (m^2/ha)
			C	F	H	MB	
1	10	C	10	5			15
2	20	C	10	10			20
4	10	F	5	5	5		15
5	10	F		25	15		40
6	10	F		25			25
7	20	F		25			25
3	10	H	10		10		20
8	10	H			3		3
9	10	H			30		30
Total	110		35	95	63	0	193
Mean			3.89	10.556	7	0	21.44

Table 24. The Phase II ground sampled polygons.

Poly #	Polygon area (ha)	Leading species	Basal area by species (m^2/ha)				Polygon basal area (m^2/ha)
			C	F	H	MB	
1	10	F		10			10
2	20	F	10	20	2		32
3	10	C	12	3	12		27
4	10	H	8		10		18
5	10	F		20	20		40
6	10	F		25	5		30
7	20	F	5	15			20
8	10	MB				15	15
9	10	H		3	15		18
Total	110		35	96	64	15	210
Mean			3.89	10.7	7.11	1.67	23.3

Table 25. The population (Phase I polygons).

Poly #	Polygon area (ha)	Leading species	Basal area by leading species (m^2/ha)				Polygon basal area (m^2/ha)
			C	F	H	MB	
1	10	C	10	5	0	0	15
2	20	C	10	10	0	0	20
3	10	H	10	0	10	0	20
4	10	F	5	5	5	0	15
5	10	F	0	25	15	0	40
6	10	F	0	25	0	0	25
7	20	F	0	25	0	0	25
8	10	H	0	0	3	0	3
9	10	H	0	0	30	0	30
10	10	H	10	0	30	0	40
11	10	C	10	5	0	0	15
12	20	F	10	10	0	0	20
13	10	H	10	0	10	0	20
14	10	F	5	5	5	0	15
15	10	F	0	25	15	0	40
16	10	F	0	25	0	0	25
17	20	F	0	25	0	0	25
18	10	H	0	0	3	0	3
19	10	H	0	0	30	0	30
20	10	H	10	0	30	0	40
21	50	F	10	10	10	0	30
Total (weighted by area)	290		1600	3100	2360	0	7060
Mean (weighted by area)			5.52	10.7	8.14	0	24.3

10.2 The algorithms

10.2.1 ADJ_{LS}

The adjustment algorithm as outlined in Section 4.1 and Figure 1 was implemented. First the adjustment was calculated by leading species group.

$$\hat{\bar{Y}}_h = \bar{y}_h + b_h(\bar{X}_h - \bar{x}_h)$$

Table 26. The adjustment statistics by stratum

Stratum	Number of polygons	Slope	Phase II mean $\bar{y}_h$	Phase I popn $\bar{X}_h$	Phase I sample $\bar{x}_h$	Adjusted popn $\hat{\bar{Y}}_h$	Ratio r_h
C	2	4.40	21	17.50	17.5	21.0	1.20
F	4	0.89	27	26.47	26.25	24.2	1.03
H	3	0.17	20	23.25	17.7	21.0	0.90

Then the adjustment ratio was calculated as follows.

$$r_h = \frac{\hat{\bar{Y}}_h}{\bar{X}_h}$$

Then the Phase I polygons were adjusted using the ratio adjustments by leading species stratum.

Table 27. The Phase I adjusted attributes.

Poly #	Leading Species	Polygon area (ha)	Ratio	Basal area by species (m^2/ha)				Adjusted basal area (m^2/ha)	Total basal area (m^2)
				C	F	H	MB		
1	C	10	1.20	12	6	0	0	18.0	180.0
2	C	20	1.20	12	12	0	0	24.0	480.0
3	H	10	0.90	9.02	0	9.02	0	18.0	180.4
4	F	10	1.03	5.14	5.14	5.14	0	15.4	154.1
5	F	10	1.03	0	25.7	15.4	0	41.1	411.0
6	F	10	1.03	0	25.7	0	0	25.7	256.9
7	F	20	1.03	0	25.7	0	0	25.7	513.7
8	H	10	0.90	0	0	2.71	0	2.7	27.1
9	H	10	0.90	0	0	27.1	0	27.1	270.6
10	H	10	0.90	9.02	0	27.1	0	36.1	360.8
11	C	10	1.20	12	6	0	0	18.0	180.0
12	F	20	1.03	10.3	10.3	0	0	20.5	411.0
13	H	10	0.90	9.02	0	9.02	0	18.0	180.4
14	F	10	1.03	5.14	5.14	5.14	0	15.4	154.1
15	F	10	1.03	0	25.7	15.4	0	41.1	411.0
16	F	10	1.03	0	25.7	0	0	25.7	256.9
17	F	20	1.03	0	25.7	0	0	25.7	513.7
18	H	10	0.90	0	0	2.71	0	2.7	27.1
19	H	10	0.90	0	0	27.1	0	27.1	270.6
20	H	10	0.90	9.02	0	27.1	0	36.1	360.8
21	F	50	1.03	10.3	10.3	10.3	0	30.8	1541.2
Total (weighted by area)				1663	3237	2242	0		7141.4
Mean (weighted by area)				5.73	11.20	7.73	0		24.6

10.2.2 ADJ_{CONST}

The adjustment algorithm as outlined in section 4.2 and Figure 2 was implemented. First the population (all species combined) adjustment was calculated.

$$\hat{\bar{Y}} = \bar{y} + b(\bar{X} - \bar{x})$$

$$25.105 = 23.33 + 0.611 * (24.34 - 21.44)$$

Then the strata ratio of means estimators were calculated.

$$r_h = \frac{\bar{y}_h}{\bar{x}_h}$$

Then the adjusted population total was calculated.

$$\hat{\bar{Y}}_h = r_h \cdot \bar{X}_h$$

Then the correction factor was calculated.

$$k = \frac{\hat{Y}}{\sum_h \hat{Y}_h}$$

Table 28. The adjustment statistics are given by stratum.

Stratum	Number of polygons	Area (ha)	ratio r_h	Phase II mean $\bar{y}_h$	Phase I popn $\bar{X}_h$	Phase I sample $\bar{x}_h$	Adjusted popn $\hat{\bar{Y}}_h$	k	$\hat{\bar{Y}}_h \cdot k$
C	2	40	1.20	21	17.50	17.5	21.00	0.961	20.19
F	4	170	1.03	27	26.47	26.25	27.23	0.961	26.17
H	3	80	1.13	20	23.25	17.7	26.32	0.961	25.30

Finally, the adjusted polygon basal areas were calculated.

$$\hat{y}_{ih} = k \cdot r_h \cdot x_{ih}$$

Table 29. The Phase I adjusted attributes.

Poly #	Leading Species	Polygon area (ha)	ratio	Basal area by species (m^2/ha)				Adjusted basal area (m^2/ha)	Total basal area (m)
				C	F	H	MB		
1	C	10	1.20	11.5	5.77	0	0	17.30	173.02
2	C	20	1.20	11.5	11.5	0	0	23.07	461.39
3	H	10	1.13	10.9	0	10.9	0	21.76	217.64
4	F	10	1.03	4.94	4.94	4.94	0	14.83	148.30
5	F	10	1.03	0	24.7	14.8	0	39.55	395.48
6	F	10	1.03	0	24.7	0	0	24.72	247.17
7	F	20	1.03	0	24.7	0	0	24.72	494.34
8	H	10	1.13	0	0	3.26	0	3.26	32.65
9	H	10	1.13	0	0	32.6	0	32.65	326.45
10	H	10	1.13	10.9	0	32.6	0	43.53	435.27
11	C	10	1.20	11.5	5.77	0	0	17.30	173.02
12	F	20	1.03	9.89	9.89	0	0	19.77	395.48
13	H	10	1.13	10.9	0	10.9	0	21.76	217.64
14	F	10	1.03	4.94	4.94	4.94	0	14.83	148.30
15	F	10	1.03	0	24.7	14.8	0	39.55	395.48
16	F	10	1.03	0	24.7	0	0	24.72	247.17
17	F	20	1.03	0	24.7	0	0	24.72	494.34
18	H	10	1.13	0	0	3.26	0	3.26	32.65
19	H	10	1.13	0	0	32.6	0	32.65	326.45
20	H	10	1.13	10.9	0	32.6	0	43.53	435.27
21	F	50	1.03	9.89	9.89	9.89	0	29.66	1483.03
Weighted total				1688	3114	2479	0		7280.55
Weighted average				5.82	10.7	8.55	0		25.11

10.2.3 ADJ_{SQB}

First the population (all species combined) adjustment was calculated.

$$\hat{\bar{Y}} = \bar{y} + b(\bar{X} - \bar{x})$$

$$25.105 = 23.33 + 0.611 * (24.34 - 21.44)$$

Then the species group ratio of means estimators were calculated.

$$r_s = \frac{\bar{y}_s}{\bar{x}_s}$$

Then the population averages, by species group s, were calculated.

$$\hat{\bar{Y}}_s = r_s \cdot \bar{X}_s$$

Then the correction factor was calculated.

$$k = \frac{\hat{Y}}{\sum_s \hat{Y}_s}$$

Table 30. The adjustment statistics.

Species group	Number of polygons	Area (ha)	ratio r_h	Phase II mean $\bar{y}_h$	Phase I popn $\bar{X}_h$	Phase I sample $\bar{x}_h$	Adjusted popn $\hat{\bar{Y}}_h$	k	$\hat{\bar{Y}}_h \cdot k$
Other	9	290	1.43	5.56	5.52	3.89	7.88	0.932	7.34
F	9	290	1.01	10.67	10.69	10.56	10.80	0.932	10.07
H	9	290	1.02	7.11	8.14	7.00	8.27	0.932	7.71
ALL	9	290		23.33	24.34	21.44	26.95		25.11

Then the species group basal areas within a polygon were adjusted.

Table 31. The Phase I adjusted attributes.

Poly #	Leading Species	Area (ha)	Basal area by species (m^2/ha)				Adjusted basal area (m^2/ha)	Total basal area (m^2)
			C	F	H	MB		
1	C	10	13.3	4.71	0	0	18.01	180.14
2	C	20	13.3	9.41	0	0	22.72	454.41
3	H	10	13.3	0	9.46	0	22.77	227.70
4	F	10	6.65	4.71	4.73	0	16.09	160.92
5	F	10	0	23.5	14.2	0	37.73	377.28
6	F	10	0	23.5	0	0	23.53	235.33
7	F	20	0	23.5	0	0	23.53	470.66
8	H	10	0	0	2.84	0	2.84	28.39
9	H	10	0	0	28.4	0	28.39	283.89
10	H	10	13.3	0	28.4	0	41.70	416.96
11	C	10	13.3	4.71	0	0	18.01	180.14
12	F	20	13.3	9.41	0	0	22.72	454.41
13	H	10	13.3	0	9.46	0	22.77	227.70
14	F	10	6.65	4.71	4.73	0	16.09	160.92
15	F	10	0	23.5	14.2	0	37.73	377.28
16	F	10	0	23.5	0	0	23.53	235.33
17	F	20	0	23.5	0	0	23.53	470.66
18	H	10	0	0	2.84	0	2.84	28.39
19	H	10	0	0	28.4	0	28.39	283.89
20	H	10	13.3	0	28.4	0	41.70	416.96
21	F	50	13.3	9.41	9.46	0	32.18	1609.18
Weighted total			2129	2918	2233	0		7280.55
Weighted average			7.34	10.1	7.7	0		25.11

11. Appendix B: VRI adjustment attributes list (from B.C. 1999b)

Attribute	Layer[2]	To be Adjusted	Phase I/II Link	Phase I	Phase II	Other	VDYP 7	TSR	Other	Attribute Category
1) Percent Canopy for Species Comp (SP0/SP1)	Main	YES	YES	Y	Y		Y			1
a) SP0 (1)	Main	YES	YES	Y	Y		Y			1
b) SP0 (2)	Main	YES	YES	Y	Y		Y			1
c) SP0 (3)	Main	YES	YES	Y	Y		Y			1
d) SP0 (4)	Main	YES	YES	Y	Y		Y			1
e) SP0 (5)	Main	YES	YES	Y	Y		Y			1
f) SP0 (6)	Main	YES	YES	Y	Y		Y			1
g) SP0 (7)	Main	YES	YES	Y	Y		Y			1
h) SP0 (8)	Main	YES	YES	Y	Y		Y			1
i) SP0 (9)	Main	YES	YES	Y	Y		Y			1
j) SP0 (10)	Main	YES	YES	Y	Y		Y			1
k) SP0 (11)	Main	YES	YES	Y	Y		Y			1
l) SP0 (12)	Main	YES	YES	Y	Y		Y			1
m) SP0 (13)	Main	YES	YES	Y	Y		Y			1
n) SP0 (14)	Main	YES	YES	Y	Y		Y			1
o) SP0 (15)	Main	YES	YES	Y	Y		Y			1
2) Species Composition:										
a) Sp1	Main	NO	YES	Y	Y		Y	Y	TEM, map	1
b) Sp1 %	Main	NO	YES	Y	Y		Y	Y	TEM, map	1
c) Sp2	Main	NO	YES	Y	Y		Y	Y	TEM, map	1
d) Sp2 %	Main	NO	YES	Y	Y		Y	Y	TEM, map	1
e) Sp3	Main	NO	YES	Y	Y		Y	Y	TEM, map	1
f) Sp3%	Main	NO	YES	Y	Y		Y	Y	TEM, map	1
g) Sp4	Main	NO	YES	Y	Y		Y	Y	TEM, map	1
h) Sp4%	Main	NO	YES	Y	Y		Y	Y	TEM, map	1
i) Sp5	Main	NO	YES	Y	Y		Y	Y	TEM, map	1
j) Sp5%	Main	NO	YES	Y	Y		Y	Y	TEM, map	1
k) Sp6	Main	NO	YES	Y	Y		Y	Y	TEM, map	1
l) Sp6%	Main	NO	YES	Y	Y		Y	Y	TEM, map	1
3) Age of Leading Species										
a) Total Age	Main	YES	YES	Y	Y		Y	Y	TEM, map	1
b) Bh Age	Main	No	No	N	N	calculated	Y	Y	TEM, map	1
c) Years to bh	Main	No	No	N	N	calculated	Y	Y	TEM, map	1
4) Age of 2nd species										
a) Total age	Main	YES	YES	Y	Y		Y	Y	TEM, map	1
5) Height of Lead species (site Ht)	Main	YES	YES	Y	Y		Y	Y	TEM, map	1
6) Lorey Ht (by SP0)										
a) Lorey Ht for Sp0(1)	Main	YES	No	N	Y	averaged	Y	N		2
b) Lorey Ht for Sp0(2)	Main	YES	No	N	Y	averaged	Y	N		2
c) Lorey Ht for Sp0(3)	Main	YES	No	N	Y	averaged	Y	N		2
d) Lorey Ht for Sp0(4)	Main	YES	No	N	Y	averaged	Y	N		2
e) Lorey Ht for Sp0(5)	Main	YES	No	N	Y	averaged	Y	N		2
f) Ht of 2nd species	Main	YES	No	Y	N		Y	Y	TEM, map	1
8) Site Index (lead Sp)	Main	No	No	No	No	calculated	Y	Y		1

9) Site Index (2nd Sp)	Main	No	No	No	No	calculated	Y	Y		1
Attribute	Layer*	To be Adjusted	Phase III Link	Phase I	Phase II	Other	VDYP 7	TSR	Other	Attribute Category
RESIDUAL LAYER (Over-story or residual)										
10) Species Composition										
a) Sp1	Resid	NO	NO	Y	N	N/A	Y	Y	TEM, map	1
b) Sp1 %	Resid	NO	NO	Y	N	N/A	Y	Y	TEM, map	1
d) Sp2	Resid	NO	NO	Y	N	N/A	Y	Y	TEM, map	1
e) Sp2 %	Resid	NO	NO	Y	N	N/A	Y	Y	TEM, map	1
f) Sp3	Resid	NO	NO	Y	N	N/A	Y	Y	TEM, map	1
g) Sp3%	Resid	NO	NO	Y	N	N/A	Y	Y	TEM, map	1
h) Sp4	Resid	NO	NO	Y	N	N/A	Y	Y	TEM, map	1
i) Sp4%	Resid	NO	NO	Y	N	N/A	Y	Y	TEM, map	1
j) Sp5	Resid	NO	NO	Y	N	N/A	Y	Y	TEM, map	1
k) Sp5%	Resid	NO	NO	Y	N	N/A	Y	Y	TEM, map	1
l) Sp6	Resid	NO	NO	Y	N	N/A	Y	Y	TEM, map	1
m) Sp6%	Resid	NO	NO	Y	N	N/A	Y	Y	TEM, map	1
11) Age of lead Species	Resid	NO	NO	Y	N	N/A	Y	Y		1
12) Age of 2nd Sp	Resid	NO	NO	Y	N	N/A	Y	Y		1
13) Height of lead Species	Resid	NO	NO	Y	N	N/A	Y	Y		1
14) Ht of 2nd Sp	Resid	NO	NO	Y	N	N/A	Y	Y		1
* Items 10 to 14 are repeated for each secondary layer (adjusted values for secondary layers are derived from main)										
15) Basal area (broken out by species)										
a) 4.0 cm	Main	YES	NO	Y	Y		Y	Y		1
b) 7.5 cm	Main	YES	NO	N	Y	surrogate	Y	Y		3
c) 12.5 cm	Main	YES	YES(1)	N	Y	surrogate	Y	Y		3
d) 17.5 cm	Main	YES	YES	N	Y	surrogate	Y	Y		3
16) # of trees/ha (broken out by species)										
a) 4.0cm	Main	YES	NO	Y	Y		Y	Y		1
b) 7.5cm	Main	YES	NO	N	Y	surrogate	Y	Y		3
c) 12.5cm	Main	YES	YES(1)	N	Y	surrogate	Y	Y		3
d) 17.5cm	Main	YES	YES	N	Y	surrogate	Y	Y		3
17) Quadratic Dbh (broken out by species)										
a) 4.0 cm	Main	NO	NO	N	N	surrogate	N	Y		2
b) 7.5cm	Main	NO	NO	N	N	surrogate	N	Y		2
c) 12.5cm	Main	NO	NO	N	N	surrogate	N	Y		2
d) 17.5cm	Main	NO	NO	N	N	surrogate	N	Y		2
18) Volume (Category 2 broken out by species)										
a) gross whole stem (4.0cm)	Main	YES	NO	N	Y	surrogate	Y	N		1

11. Appendix B: VRI adjustment attributes list (from B.C. 1999b)

Attribute	Layer	To be Adjusted	Phase I/II Link	Phase I	Phase II	Other	VDYP 7	TSR	Other	Attribute Category
1) Genus groups for Sp Comp (SP0(n))	Main	YES(?)	YES	Y	Y		Y			2
a) SP0(1)	Main	YES(?)	YES	Y	Y		Y			2
b) SP0(2)	Main	YES(?)	YES	Y	Y		Y			2
c) SP0(3)	Main	YES(?)	YES	Y	Y		Y			2
d) SP0(4)	Main	YES(?)	YES	Y	Y		Y			2
e) SP0 (5)	Main	YES(?)	YES	Y	Y		Y			2
f) SP0 (6)	Main	YES(?)	YES	Y	Y		Y			2
g) SP0(1) %	Main	YES(?)	YES	Y	Y		Y			2
h) SP0(2) %	Main	YES(?)	YES	Y	Y		Y			2
i) SP0(3) %	Main	YES(?)	YES	Y	Y		Y			2
j) SP0(4) %	Main	YES(?)	YES	Y	Y		Y			2
k) SP0(5) %	Main	YES(?)	YES	Y	Y		Y			2
l) SP0(6) %	Main	YES(?)	YES	Y	Y		Y			2
2) Species Composition:										
a) Sp1	Main	NO	YES	Y	Y		Y	Y	TEM, map	1
b) Sp1 %	Main	NO	YES	Y	Y		Y	Y	TEM, map	1
c) Sp2	Main	NO	YES	Y	Y		Y	Y	TEM, map	1
d) Sp2 %	Main	NO	YES	Y	Y		Y	Y	TEM, map	1
e) Sp3	Main	NO	YES	Y	Y		Y	Y	TEM, map	1
f) Sp3%	Main	NO	YES	Y	Y		Y	Y	TEM, map	1
g) Sp4	Main	NO	YES	Y	Y		Y	Y	TEM, map	1
h) Sp4%	Main	NO	YES	Y	Y		Y	Y	TEM, map	1
i) Sp5	Main	NO	YES	Y	Y		Y	Y	TEM, map	1
j) Sp5%	Main	NO	YES	Y	Y		Y	Y	TEM, map	1
k) Sp6	Main	NO	YES	Y	Y		Y	Y	TEM, map	1
l) Sp6%	Main	NO	YES	Y	Y		Y	Y	TEM, map	1
3) Age of Leading Species										
a) Total Age	Main	YES	YES	Y	Y		Y	Y	TEM, map	1
b) Bh Age	Main	No	No	N	N	calculated	Y	Y	TEM, map	1
c) Years to bh	Main	No	No	N	N	calculated	Y	Y	TEM, map	1
4) Age of 2nd species										
a) Total age	Main	YES	YES	Y	Y		Y	Y	TEM, map	1
5) Height of Lead species (site Ht)	Main	YES	YES	Y	Y		Y	Y	TEM, map	1
6) Lorey Ht (by SP0)										
a) Lorey Ht for Sp0(1)	Main	YES	No	N	Y	surrogate	Y	N		2
b) Lorey Ht for Sp0(2)	Main	YES	No	N	Y	surrogate	Y	N		2
c) Lorey Ht for Sp0(3)	Main	YES	No	N	Y	surrogate	Y	N		2
d) Lorey Ht for Sp0(4)	Main	YES	No	N	Y	surrogate	Y	N		2
e) Lorey Ht for Sp0(5)	Main	YES	No	N	Y	surrogate	Y	N		2
f) Lorey Ht for Sp0(6)	Main	YES	No	N	Y	surrogate	Y	N		2
7) Ht of 2nd Species	Main	YES	No	Y	N		Y	Y	TEM, map	1
8) Site Index (lead Sp)	Main	No	No	No	No	calculated	Y	Y		1

9) Site Index (2nd Sp)	Main	No	No	No	No	calculated	Y	Y		1
Attribute	Layer*	To be Adjusted	Phase I/II Link	Phase I	Phase II	Other	VDYP 7	TSR	Other	Attribute Category
RESIDUAL LAYER (Over-story or residual)										
10) Species Composition										
a) Sp1	Resid	NO	NO	Y	N	N/A	Y	Y	TEM, map	1
b) Sp1 %	Resid	NO	NO	Y	N	N/A	Y	Y	TEM, map	1
d) Sp2	Resid	NO	NO	Y	N	N/A	Y	Y	TEM, map	1
e) Sp2 %	Resid	NO	NO	Y	N	N/A	Y	Y	TEM, map	1
f) Sp3	Resid	NO	NO	Y	N	N/A	Y	Y	TEM, map	1
g) Sp3%	Resid	NO	NO	Y	N	N/A	Y	Y	TEM, map	1
h) Sp4	Resid	NO	NO	Y	N	N/A	Y	Y	TEM, map	1
i) Sp4%	Resid	NO	NO	Y	N	N/A	Y	Y	TEM, map	1
j) Sp5	Resid	NO	NO	Y	N	N/A	Y	Y	TEM, map	1
k) Sp5%	Resid	NO	NO	Y	N	N/A	Y	Y	TEM, map	1
l) Sp6	Resid	NO	NO	Y	N	N/A	Y	Y	TEM, map	1
m) Sp6%	Resid	NO	NO	Y	N	N/A	Y	Y	TEM, map	1
11) Age of lead Species	Resid	NO	NO	Y	N	N/A	Y	Y		1
12) Age of 2nd Sp	Resid	NO	NO	Y	N	N/A	Y	Y		1
13) Height of lead Species	Resid	NO	NO	Y	N	N/A	Y	Y		1
14) Ht of 2nd Sp	Resid	NO	NO	Y	N	N/A	Y	Y		1
* Items 10 to 14 are repeated for each secondary layer (adjusted values for secondary layers are derived from main)										
15) Basal area (broken out by species)										
a) 4.0 cm	Main	YES	NO	Y	Y		Y	Y		1
b) 7.5 cm	Main	YES	NO	N	Y	surrogate	Y	Y		2
c) 12.5 cm	Main	YES	YES(?)	N	Y	surrogate	Y	Y		2
d) 17.5 cm	Main	YES	YES	N	Y	surrogate	Y	Y		2
16) # of trees/ha (broken out by species)										
a) 4.0cm	Main	YES	NO	Y	Y		Y	Y		1
b) 7.5cm	Main	YES	NO	N	Y	surrogate	Y	Y		2
c) 12.5cm	Main	YES	YES(?)	N	Y	surrogate	Y	Y		2
d) 17.5cm	Main	YES	YES	N	Y	surrogate	Y	Y		2
17) Quadratic Dbh (broken out by species)										
a) 4.5 cm	Main	NO	NO	N	N	calculated	N	Y		2
b) 7.5cm	Main	NO	NO	N	N	calculated	N	Y		2
c) 12.5cm	Main	NO	NO	N	N	calculated	N	Y		2
d) 17.5cm	Main	NO	NO	N	N	calculated	N	Y		2
18) Volume (Category 2 broken out by species)										
a) gross whole stem (4.0cm)	Main	YES	NO	N	Y	surrogate	Y	N		1
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
iv) net top & stump (CU)	Main	YES	NO	N	Y	surrogate	Y	N		2

Attribute	Layer[2]	To be Adjusted	Phase I/II Link	Phase I	Phase II	Other	VDYP 7	TSR	Other	Attribute Category
b) 7.5cm (gross whole stem)	Main	YES	NO	N	Y	surrogate	Y	N		2
i) net top & stump (CU)	Main	YES	NO	N	Y	surrogate	Y	N		2
ii) net decay	Main	YES	NO	N	Y	surrogate	Y	N		2
iii) net decay & waste	Main	YES	NO	N	Y	surrogate	Y	N		2
iv) net decay waste & breakage	Main	YES	NO	N	Y	surrogate	Y	N		1
c) 12.5cm (gross whole stem)	Main									2
i) net top & stump (CU)	Main	YES	NO	N	Y	surrogate	Y	Y		2
ii) net decay	Main	YES	NO	N	Y	surrogate	Y	Y		2
iii) net decay & waste	Main	YES	NO	N	Y	surrogate	Y	Y		2
iv) net decay waste & breakage	Main	YES	YES	N	Y	surrogate	Y	Y		1
d) 17.5cm (gross whole stem)	Main	YES	NO	N	Y	surrogate	Y	Y		2
i) net top & stump (CU)	Main	YES	NO	N	Y	surrogate	Y	Y		2
ii) net decay	Main	YES	NO	N	Y	surrogate	Y	Y		2
iii) net decay & waste	Main	YES	NO	N	Y	surrogate	Y	Y		2
iv) net decay waste & breakage	Main	YES	YES	N	Y	surrogate	Y	Y		1
iv) net decay waste & breakage	Main	YES		N	Y	surrogate	Y	N		1
19) Dead tree volume										
a) gross whole stem (4.0cm)	Main	YES		N	Y	surrogate	N	Y	Eco/biomass	3
b) gross whole stem (7.5cm)	Main	YES		N	Y	surrogate	N	Y	Eco/biomass	3
c) 12.5cm	Main	YES		N	Y	surrogate	N	Y	Eco/biomass	3
i) net decay	Main	YES		N	Y	surrogate	N	Y	Eco/biomass	3
ii) net decay, waste & [illegible]	Main	YES		N	Y	surrogate	N	Y	Eco/biomass	3

Attribute	Layer[2]	To be Adjusted	Phase I/II Link	Phase I	Phase II	Other	VDYP 7	TSR	Other	Attribute Category
20) Coarse woody debris										
a) decay class 1	Main	YES	NO	N	Y	surrogate	N	Y	Eco/biomass	3
b) decay class 2	Main	YES	NO	N	Y	surrogate	N	Y	Eco/biomass	3
c) decay class 3	Main	YES	NO	N	Y	surrogate	N	N	Eco/biomass	3
d) decay class 4	Main	YES	NO	N	Y	surrogate	N	N	Eco/biomass	3
e) decay class 5	Main	YES	NO	N	Y	surrogate	N	N	Eco/biomass	3
21) Stump volume (4.0cm+)		YES	NO	N	Y	surrogate	N	Y	Eco/biomass	3
22) # of snags/ha (4.0cm+)		YES	YES	Y	Y		N	Y	Eco/biomass	1
23) Shrub Ht		YES	YES	Y	Y		N	Y		1
24) BEC based Site Index		NO	NO	Y	N		N	Y		3
25) Biogeoclimatic Zone (BEC)										
a) zone		NO					Y			1
b) sub-zone		NO					Y			1
26) Mapsheet		NO					Y			1
27) Polygon number		NO					Y			1
28) Reference year		NO					Y			1
29) Projection year		NO					Y			1
LEGACY ATTRIBUTES										
30) Stocking class(^3)		NO	NO	N	N		?	N	transient	3
31) Crown closure		NO	NO	Y	N		?	Y	Eco	3
32) Forest Inventory Zone (FIZ)									compilation	3

The following explanations reffer to comments embedded with the spread-sheets:

(^1) - 1. This list does not include projected attributes and other outputs generated by VDYP7 (, i.e., compatibility attributes).

(^2) - 2. In some cases several non-overstory layers will be aggregated to create the G_layer for adjustment. After adjustment, the information will be disaggregated back to the layers.

(^3) - 3. Some attributes from FIP files will be carried over to the new data base until new photo interpretation is acquired.

Attribute	Layer[3]	To be Adjusted	Phase I/II Link	Phase I	Phase II	Other	VDYP 7	TSR	Other	Attribute Category
[illegible]	[illegible]	YES	NO	N	Y	surrogate	Y	N		1
ii) net bkp & stump (CU)	Main	YES	NO	N	Y	surrogate	Y	N		2
iii) net decay	Main	YES	NO	N	Y	surrogate	Y	N		2
iii) net decay & [illegible]	Main	YES	NO	N	Y	surrogate	Y	N		2
iv) net decay waste & breakage	Main	YES	NO	N	Y	surrogate	Y	N		1
[illegible]	Main									2
ii) net bkp & stump (CU)	Main	YES	NO	N	Y	surrogate	Y	Y		2
iii) net decay	Main	YES	NO	N	Y	surrogate	Y	Y		2
iii) net decay & [illegible]	Main	YES	NO	N	Y	surrogate	Y	Y		2
iv) net decay waste & breakage	Main	YES	YES	N	Y	surrogate	Y	Y		1
[illegible]	Main	YES	NO	N	Y	surrogate	Y	Y		2
ii) net bkp & stump (CU)	Main	YES	NO	N	Y	surrogate	Y	Y		2
iii) net decay	Main	YES	NO	N	Y	surrogate	Y	Y		2
iii) net decay & [illegible]	Main	YES	NO	N	Y	surrogate	Y	Y		2
iv) net decay waste & breakage	Main	YES	YES	N	Y	surrogate	Y	Y		1
iv) net decay waste & breakage	Main	YES		N	Y	surrogate	Y	N		1
19) Dead tree volume										
[illegible]										

Attribute	Layer	To be Adjusted	Phase I/II Link	Phase I	Phase II	Other	VDYP 7	TSR	Other	Attribute Category
20) Coarse woody debris										
22) # of snags/ha (4.0cm+)		YES	YES	Y	Y		N	Y	Eco/biomass	1
23) Shrub Ht		YES	YES	Y	Y		N	Y		1
25) Biogeoclimatic Zone (BEC)										
a) zone		NO					Y			1
b) sub-zone		NO					Y			1
26) Mapsheet		NO					Y			1
27) Polygon number		NO					Y			1
28) Reference year		NO					Y			1
29) Projection year		NO					Y			1
LEGACY ATTRIBUTES										

The following explanations refer to comments embedded with the spread-sheets:

(^1) - 1. This list does not include projected attributes and other outputs generated by VDYP7 (, i.e., compatibility attributes).

(^2) - 2. In some cases several non-overstory layers will be aggregated to create the G_layer for adjustment. After adjustment, the information will be disaggregated back to the layers.

(^3) - 3. Some attributes from FIP files will be carried over to the new data base until new photo interpretation is acquired.